AD-A012 635

AN OPTIMAL SYSTEM FOR TRACKING A SINUSOID OF TIME-VARYING FREQUENCY BURIED IN ADDITIVE NOISE

C. Nicholas Pryor

Naval Surface Weapons Center
Silver Spring, Maryland

2 June 1975

NSWC

TECHNICAL REPORT

WHITE OAK LABORATORY
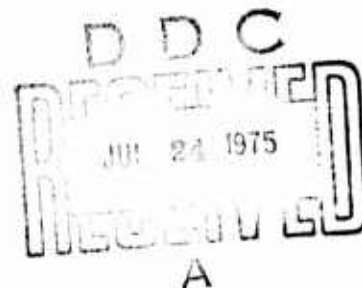
AN OPTIMAL SYSTEM FOR TRACKING A SINUSOID OF TIME-VARYING FREQUENCY
BURIED IN ADDITIVE NOISE

BY
C. Nicholas Pryor   2 JUNE 1975

NAVAL SURFACE WEAPONS CENTER
WHITE OAK LABORATORY
SILVER SPRING, MD. 20910

DDC
JUL 24 1975
A

**NAVAL SURFACE WEAPONS CENTER
WHITE OAK, SILVER SPRING, MARYLAND 20910**

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NSWC/WOL/TR 75-74	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Optimal System For Tracking A Sinusoid Of Time-Varying Frequency Buried In Additive Noise		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) C. Nicholas Pryor		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Surface Weapons Center White Oak, Silver Spring, MD 20910		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS AIR WF11-121-711/A370-370A
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE 2 June 1975
		13. NUMBER OF PAGES 74
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Signal Processing
Phase Lock Loop
Frequency Modulation
Kalman Filtering

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

It is often necessary to provide a continuous estimate of the frequency of a sinusoid of randomly varying frequency, through a channel containing noise and fading. This problem is modeled and the optimum estimator derived using continuous Kalman filter theory. The optimum system is shown to be equivalent to a phase lock loop, with some special features added

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |

PRICES SUBJECT TO CHANGE

i

when Rayleigh fading is considered. The steady state solution of the Kalman error matrix is found. This establishes all parameters of the optimum processor and provides a measure of its performance as a function of input signal and noise characteristics.

2 June 1975

This report derives the optimum system for tracking the
frequency of a signal through a noisy and fading propagation
channel. This problem is of considerable interest in both
underwater acoustics and in high-frequency radio propagation.
The optimum system is shown to be similar to the commonly
used phase lock loop, and optimum parameters and performance
figures are derived.

The majority of the work reported was performed while the author
was a member of the Signal Processing Division of the White Oak
Laboratory, under task Air WF11-121-711/A370-370A. Many of the
results were presented at the USAG Minisymposium on Automatic
Signal Processing at New London in June 1974.

C. Nicholas Pryor

C. NICHOLAS PRYOR
By direction

CONTENTS

ILLUSTRATIONS

ILLUSTRATIONS (Continued)

INTRODUCTION

In certain applications it is desirable to perform
a continuous measurement of the frequency of a sine wave
when the sine wave is immersed in a random noise background.
This is a common application of systems such as phase lock
loops, as well as other forms of frequency modulation
detectors. If the frequency of the sine wave is very stable
over a long period of time, then its frequency can be
measured with arbitrary accuracy. However, if the
frequency fluctuates in some random manner with time
and the goal is to track these frequency fluctuations in the
presence of noise, then an optimum system can be derived for
estimating the frequency versus time and a lower bound can
be placed on the mean squared tracking error. This is the
topic to be explored in this report, with the goal of providing
insight into the design and performance of frequency trackers.
While the analytical approach used is based on continuous
Kalman filtering theory* the resulting system closely resembles
a conventional phase lock loop. Thus the results can be applied
to the design of an adaptive phase lock loop tracking system.

*The notation used in this report follows that in Reference (1)

ANALYTICAL MODEL

If the signal to be tracked is a sine wave with instantaneous amplitude A and radian frequency ω, it may be represented at any instant as A cos θ where $\theta = \int \omega$ dt. The instantaneous state of the input signal process may be represented by a three-dimensional vector

$$X] = \begin{bmatrix} \theta \\ \omega \\ A \end{bmatrix} = \begin{bmatrix} \text{phase in radians} \\ \text{frequency in rad/sec} \\ \text{amplitude in volts} \end{bmatrix} \qquad (1)$$

The frequency parameter ω is the state variable which we wish to measure, while θ and A are required to define the measure

$$h(X) = h(\theta,\omega,A) = A \cos\theta \qquad (2)$$

from which we must infer the state. If a measurement z is taken of the input signal plus noise process, then z = h(X) + v where v is a zero-mean random process representing the contamination of the measurement due to noise. If the noise is a white noise process with a (double sided frequency) power density of N volts^2/Hz, a noise autocorrelation matrix R can be defined as

$$\overline{v(t)\ v(t+\tau)^T} = [R]\delta(\tau) = N\ \delta(\tau) \qquad (3)$$

where R becomes a 1x1 matrix equal to N.

The evolution of the state of the input process with time may be described by the set of differential equations

$$\dot{X}] = [F] \cdot X] + [G] \cdot w] \tag{4}$$

where $\dot{X}]$ is the first derivative of the state $X]$, the matrix

$$[F] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\alpha & 0 \\ 0 & 0 & -\beta \end{bmatrix} \tag{5}$$

defines the dynamic behavior of the input process, $w]$ is a random white noise vector which produces the random fluctuations in the input signal, and $[G]$ is a matrix which couples this random process to the system. The random vector $w]$ is described by its mean $w_o]$ and a covariance matrix $[Q]$. In order to provide random fluctuations in both frequency and amplitude, $w]$ is defined as a two-dimensional vector such that

$$w_o] = \begin{bmatrix} \omega_o \\ A_o \end{bmatrix} \tag{6}$$

and

$$\overline{[w(t) - w_o][w(t+\tau) - w_o]^T} = [Q]\delta(\tau) = \begin{bmatrix} Q_\omega & 0 \\ 0 & Q_A \end{bmatrix} \delta(\tau) \tag{7}$$

The matrix G is defined as

$$G = \begin{bmatrix} 0 & 0 \\ \alpha & 0 \\ 0 & \beta \end{bmatrix} \tag{8}$$

With these definitions, ω_O and A_O become the mean frequency and the mean amplitude of the sinusoidal process respectively, while Q_ω and Q_A respectively determine the sizes of independent frequency and amplitude perturbations around these means. As a result of the form of [F], these perturbations both resemble simple low-pass random processes which may be described by their autocorrelation functions

$$\phi_{\omega\omega}(\tau) = (\alpha Q_\omega/2) \exp(-\alpha|\tau|) \tag{9}$$

and

$$\phi_{AA}(\tau) = (\beta Q_A/2) \exp(-\beta|\tau|) \tag{10}$$

The rms frequency deviation from the mean ω_O is thus $\sqrt{\alpha Q_\omega/2}$ radians/sec, and the correlation time constant (or typical duration) of these fluctuations is $1/\alpha$ seconds.* Similarly the rms deviation of the amplitude from the mean A_O is $\sqrt{\beta Q_A/2}$ volts and the typical time associated with these

---

*α may also be called the bandwidth of the random frequency fluctuations, but one must be careful to distinguish this modulation "bandwidth" from the rms frequency deviation of the signal.

fluctuations is $1/\beta$ seconds. While this model is not
intended to represent any particular system, it is
sufficiently general to represent a wide range of sinusoidal
processes having some degree of both amplitude and frequency
random modulation. It contains sufficient parameters to
characterize both a magnitude and a bandwidth (or typical
period) associated with each type of fluctuation.

DERIVATION OF KALMAN FILTER

The model presented in the previous section is that of
a continuous linear system described by a set of first
order differential equations, on which a single continuous
measurement is made where the measurement is a nonlinear function
of the system state. The theory of Kalman filtering
provides a means of estimating the state of this system
as a function of time, using the noisy input measurements,
such that the mean squared error in the estimate is minimized.
The Kalman filter provides a continuous estimate $\hat{X}]$ of the current
state $X]$ of the input process plus a covariance matrix $[P]$
representing the estimated uncertainty in the estimate of $X]$.
In the present problem

$$[P] = \overline{[X - \hat{X}][X - \hat{X}]^T} = \begin{bmatrix} P_{\theta\theta} & P_{\theta\omega} & P_{\theta A} \\ P_{\theta\omega} & P_{\omega\omega} & P_{\omega A} \\ P_{\theta A} & P_{\omega A} & P_{AA} \end{bmatrix} \tag{11}$$

contains the estimated variance of the estimate of each state
variable plus the covariance between them. Reference (1) gives
the form of the continuous Kalman filter as a system obeying the
differential equations

$$\dot{\hat{X}} = F \hat{X} + G w_o + P H^T R^{-1} (z - h(\hat{X})) \tag{12}$$

and

$$\dot{P} = F P + P F^T + G Q G^T - P H^T R^{-1} H P \tag{13}$$

The H matrix in this set of equations is a linearized form of the measurement function and consists of the Jacobian matrix

$$[H] = dh/d\hat{X} \tag{14}$$

In our problem this becomes a 1x3 matrix

$$H = [\partial h/\partial\hat{\theta}, \ \partial h/\partial\hat{\omega}, \ \partial h/\partial\hat{A}]$$

$$= [-\hat{A} \sin\hat{\theta}, \ 0, \ \cos\hat{\theta}] \tag{15}$$

Thus, given an initial estimate $\hat{X}(0)$ of the state and an estimate P(0) of the initial uncertainty, these matrix equations provide a way of computing the best estimate $\hat{X}$ as a function of time as well as showing the way the uncertainty matrix evolves with time. Note that the equation for the behavior of P depends only on the several matrices defining the problem and not at all on the measurements z of the input. The differential equations for the state estimate $\hat{X}$ do use the measurement inputs and require P in their solution. Thus the general form for the Kalman filter consists of a time-varying filter which determines the $\hat{X}$ estimates from the input data, where the coefficients of this filter vary with the estimated error matrix P. However in problems such as the one studied here the system model and all noise matrices are stationary in time so that a steady-state

value will eventually be reached for the P matrix. A steady-state form of the filter can thus be found by setting $\dot{P}$ equal to zero in equation (13), solving for P, and then using this steady-state P matrix in equation (12) for the state estimation filter.

EXPANSION OF KALMAN FILTER EQUATIONS

In principle the freqeuncy tracking problem is solved at this point. A time-varying Kalman filter may be implemented directly by repetitive matrix solution of equations (12) and (13) using the matrices as defined in the model and using the residuals $(z-h(\hat{X}))$ of the input measurements as inputs to the filter. In large Kalman filters (having a large number of state and/or measurement variables) this direct matrix approach is about the only method available because analytical expansion of the deceptively simple looking matrix equations is impractical. In this rather small problem, however, it is possible to expand the indicated matrix operations analytically and gain much insight into operations actually being performed by the filter.

The first necessary step is to expand the matrix equation (13) for $\dot{P}$ to form

$$\dot{P} = F P + P F^T + G Q G^T - P H^T R^{-1} H P$$

$$= \begin{bmatrix} P_{\theta\omega} & P_{\omega\omega} & P_{\omega A} \\ -\alpha P_{\theta\omega} & -\alpha P_{\omega\omega} & -\alpha P_{\omega A} \\ -\beta P_{\theta A} & -\beta P_{\omega A} & -\beta P_{AA} \end{bmatrix} + \begin{bmatrix} P_{\theta\omega} & -\alpha P_{\theta\omega} & -\beta P_{\theta A} \\ P_{\omega\omega} & -\alpha P_{\omega\omega} & -\beta P_{\omega A} \\ P\omega_A & -\alpha P_{\omega A} & -\beta P_{AA} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \alpha^2 Q_\omega & 0 \\ 0 & 0 & \beta^2 Q_A \end{bmatrix}$$

$$- (1/N) \begin{bmatrix} -AP_{\theta\theta} \sin\theta + P_{\theta A} \cos\theta \\ -AP_{\theta\omega} \sin\theta + P_{\omega A} \cos\theta \\ -AP_{\theta A} \sin\theta + P_{AA} \cos\theta \end{bmatrix} \cdot \begin{bmatrix} -AP_{\theta\theta} \sin\theta + P_{\theta A} \cos\theta \\ -AP_{\theta\omega} \sin\theta + P_{\omega A} \cos\theta \\ -AP_{\theta A} \sin\theta + P_{AA} \cos\theta \end{bmatrix}^T \quad (16)$$

12

From this matrix equation we can extract the six scalar differential equations for the components of $\dot{P}$

$$\dot{P}_{\theta\theta} = 2P_{\theta\omega} - (1/N)(-AP_{\theta\theta}\sin\theta + P_{\theta A}\cos\theta)^2 \tag{17a}$$

$$\dot{P}_{\omega\omega} = -2\alpha P_{\omega\omega} + \alpha^2 Q_\omega - (1/N)(-AP_{\theta\omega}\sin\theta + P_{\omega A}\cos\theta)^2 \tag{17b}$$

$$\dot{P}_{AA} = -2\beta P_{AA} + \beta^2 Q_A - (1/N)(-AP_{\theta A}\sin\theta + P_{AA}\cos\theta)^2 \tag{17c}$$

$$\dot{P}_{\theta\omega} = P_{\omega\omega} - \alpha P_{\theta\omega} - (1/N)(-AP_{\theta\theta}\sin\theta + P_{\theta A}\cos\theta)(-AP_{\theta\omega}\sin\theta + P_{\omega A}\cos\theta) \tag{17d}$$

$$\dot{P}_{\theta A} = P_{\omega A} - \beta P_{\theta A} - (1/N)(-AP_{\theta\theta}\sin\theta + P_{\theta A}\cos\theta)(-AP_{\theta A}\sin\theta + P_{AA}\cos\theta) \tag{17e}$$

$$\dot{P}_{\omega A} = -(\alpha+\beta)P_{\omega A} - (1/N)(-AP_{\theta\omega}\sin\theta + P_{\omega A}\cos\theta)(-AP_{\theta A}\sin\theta + P_{AA}\cos\theta) \tag{17f}$$

Solving these six differential equations continuously and exactly is a rather awesome task; fortunately some simplifications are possible. First, there is an implicit time dependence in the coefficients of these equations since $\sin\theta$ and $\cos\theta$ both vary at the input signal frequency. This time dependence is induced by the dependence of the measurement matrix H on the signal phase θ. If we assume that fluctuations in the P matrix (that is, variations in uncertainty of the state estimate) occurring within a single period of the input sinusoid are unimportant, then time-average values of $\dot{P}$ may be used in determining the dynamics of P, where the time average is taken over a period of the input sinusoid. Using $\langle\cos^2\theta\rangle = \langle\sin^2\theta\rangle = 1/2$ and $\langle\sin\theta\cos\theta\rangle = 0$, the time-average form of equations (17) becomes

13

$$\langle \dot{P}_{\theta\theta} \rangle = 2P_{\theta\omega} - (1/2N)(A^2 P_{\theta\theta}^2 + P_{\theta A}^2) \tag{18a}$$

$$\langle \dot{P}_{\omega\omega} \rangle = -2\alpha P_{\omega\omega} + \alpha^2 Q_\omega - (1/2N)(A^2 P_{\theta\omega}^2 + P_{\omega A}^2) \tag{18b}$$

$$\langle \dot{P}_{AA} \rangle = -2\beta P_{AA} + \beta^2 Q_A - (1/2N)(A^2 P_{\theta A}^2 + P_{AA}^2) \tag{18c}$$

$$\langle \dot{P}_{\theta\omega} \rangle = P_{\omega\omega} - \alpha P_{\theta\omega} - (1/2N)(A^2 P_{\theta\theta} P_{\theta\omega} + P_{\theta A} P_{\omega A}) \tag{18d}$$

$$\langle \dot{P}_{\theta A} \rangle = P_{\omega A} - \beta P_{\theta A} - (1/2N)(A^2 P_{\theta\theta} P_{\theta A} + P_{\theta A} P_{AA}) \tag{18e}$$

$$\langle \dot{P}_{\omega A} \rangle = -(\alpha+\beta)P_{\omega A} - (1/2N)(A^2 P_{\theta\omega} P_{\theta A} + P_{\omega A} P_{AA}) \tag{18f}$$

A second simplification is now possible by recognizing that equations (18e) and (18f) are linear in $P_{\theta A}$ and $P_{\omega A}$ and that there are no other terms driving them as inputs. Their solutions are decaying exponentials so that both of these covariances decay to zero if for any reason they are initially given non-zero values. This is physically reasonable since there is no coupling mechanism in the model between amplitude and phase or frequency. Thus one would expect their estimation errors to be uncorrelated. Note that this reasoning does not apply to $P_{\theta\omega}$ since errors in frequency naturally evolve into errors in phase. However the fact that $P_{\theta A}$ and $P_{\omega A}$ can be assumed equal to zero reduces the number of equations to four and simplifies those remaining to

$$\langle \dot{P}_{\theta\theta} \rangle = 2P_{\theta\omega} - (A^2/2N)P_{\theta\theta}^2 \tag{19a}$$

$$\langle \dot{P}_{\omega\omega} \rangle = \alpha^2 Q_\omega - 2\alpha P_{\omega\omega} - (A^2/2N)P_{\theta\omega}^2 \tag{19b}$$

$$\langle \dot{P}_{AA} \rangle = \beta^2 Q_A - 2\beta P_{AA} - (1/2N)P_{AA}^2 \tag{19c}$$

$$\langle \dot{P}_{\theta\omega} \rangle = P_{\omega\omega} - \alpha P_{\theta\omega} - (A^2/2N)P_{\theta\theta} P_{\theta\omega} \tag{19d}$$

Having found the differential equations representing the behavior of the uncertainty matrix P with time, we can proceed in a similar fashion with the state estimation equation, equation (12). Substituting and multiplying the indicated matrixes in equation (12) gives

$$\dot{\hat{X}} = F\hat{X} + G w_o + P H^T R^{-1}(z - h(\hat{X}))$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\alpha & 0 \\ 0 & 0 & -\beta \end{bmatrix} \cdot \begin{bmatrix} \hat{\theta} \\ \hat{\omega} \\ \hat{A} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \alpha & 0 \\ 0 & \beta \end{bmatrix} \cdot \begin{bmatrix} \omega_o \\ A_o \end{bmatrix} + (1/N)\begin{bmatrix} -AP_{\theta\theta}\sin\theta \\ -AP_{\theta\omega}\sin\theta \\ P_{AA}\cos\theta \end{bmatrix} \cdot (z-\hat{A}\cos\hat{\theta})$$

$$(20)$$

where we have taken advantage of the fact that $P_{\theta A}$ and $P_{\omega A}$ are equal to zero. Writing these out as individual differential equations gives

$$\dot{\hat{\theta}} = \hat{\omega} - (A/N)P_{\theta\theta}\sin\theta \ (z-\hat{A}\cos\hat{\theta}) \tag{21a}$$

$$\dot{\hat{\omega}} = \alpha(\omega_o-\hat{\omega}) - (A/N)P_{\theta\omega}\sin\theta \ (z-\hat{A}\cos\hat{\theta}) \tag{21b}$$

$$\dot{\hat{A}} = \beta(A_o-A) + (1/N)P_{AA}\cos\theta \ (z-\hat{A}\cos\hat{\theta}) \tag{21c}$$

Just as in the equations for $\dot{P}$, these equations contain some terms which vary sinusoidally at the signal frequency or higher. Assuming that these short term variations are meaningless as far as estimation of the state is concerned, we can again take time averages of these equations over the sinusoidal signal period to obtain

15

$$\dot{\langle \hat{\theta} \rangle} = \hat{\omega} - (A/N)P_{\theta\theta} \langle z \sin\hat{\theta} \rangle \tag{22a}$$

$$\dot{\langle \hat{\omega} \rangle} = \alpha(\omega_o - \hat{\omega}) - (A/N) P_{\theta\omega} \langle z \sin\hat{\theta} \rangle \tag{22b}$$

$$\dot{\langle \hat{A} \rangle} = \beta(A_o - \hat{A}) + (1/N) P_{AA}[\langle z \cos\hat{\theta} \rangle - (A/2)] \tag{22c}$$

Note that once the equations are written in this form the
concept of forming the residual $(z-h(\hat{X}))$ and using it to
update the estimate vanishes and is replaced by the idea of
using the mean products of z with $\sin\hat{\theta}$ and $\cos\hat{\theta}$ as the basic
information inputs to the filter.

Equations (19) and (22) thus become the equations which must
be solved to implement the time-variable Kalman filter, where
advantage has been taken of the lack of correlation
between amplitude and frequency estimation errors and of
smoothing of the equations over the period of the input
signal. Assuming some method is available for solving the
nonlinear equations (19) to obtain P, the state estimation
equations, (22), can be implemented by using a system
such as that shown in Figure 1. Each integrator output
represents one of the state variables, while its input is the
derivative of that state variable formed according to equations (22).

Several interesting things can be learned from the form of
the system shown in Figure 1. First is that, while the
subsystem for estimating amplitude is nearly independent of that
for estimating frequency and phase, each depends on the other
for some piece of information. An estimate of A is used to
establish gain coefficients in the frequency subsystem, while the
estimate of θ is required to generate the product $\langle z \cos\hat{\theta} \rangle$
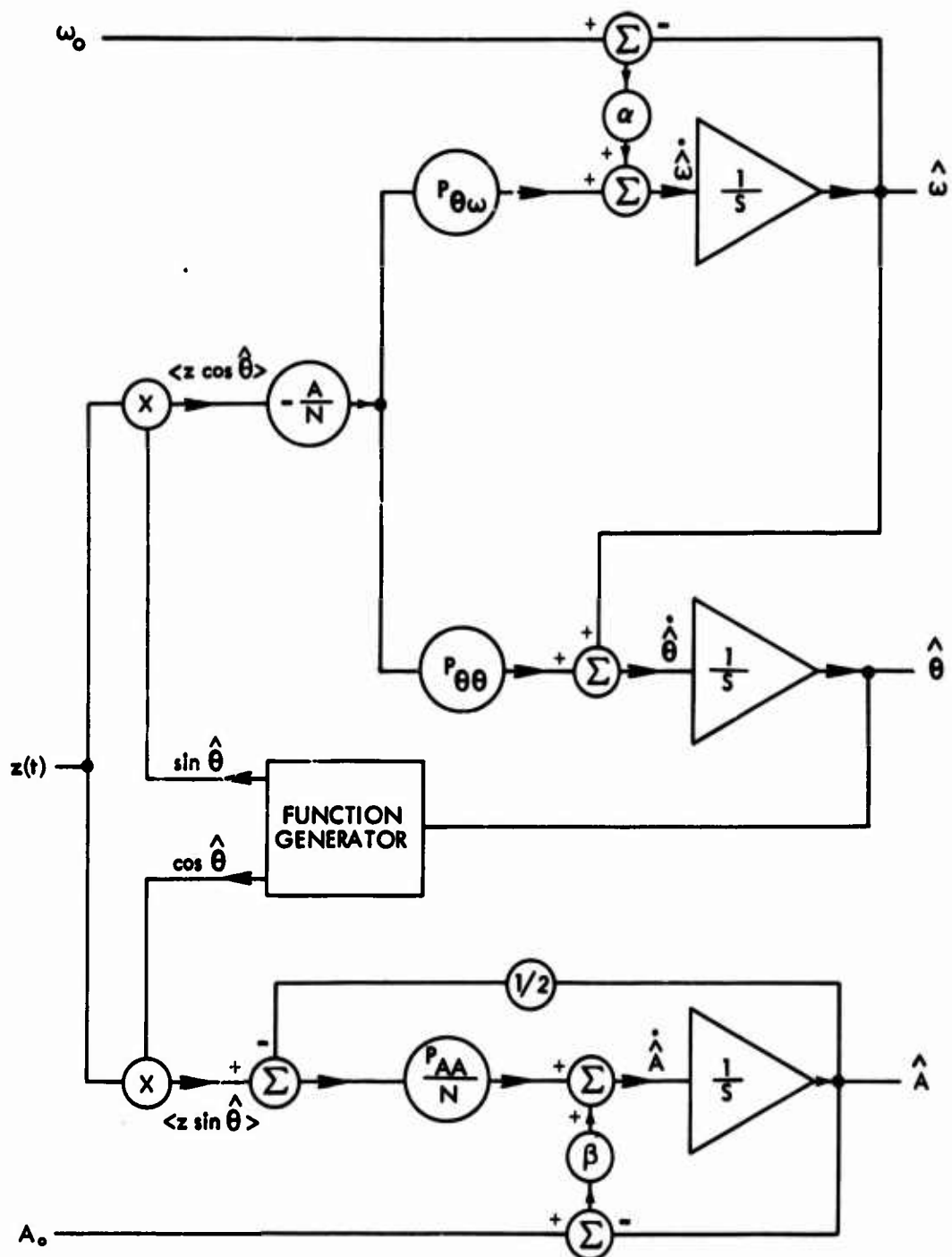
16

FIG. 1   STATE ESTIMATION PORTION OF KALMAN FILTER

used as the input to the amplitude subsystem. Values of $P_{\theta\theta}$, $P_{\theta\omega}$, and P_{AA} are also used by the state estimation system to establish gain coefficients. Both optimal subsystems have forms which are commonly used in systems of this type. The subsystem for estimating A is really a simple low-pass filter (or exponential averager) with a 3 db bandwidth of $\beta + P_{AA}/2N$ radians/second (or a time constant equal to the inverse of this bandwidth). It is driven by two inputs, one of which is the presumably known average signal level A_0 and the other is the measured "in phase" component of the input process $\langle z \cos\hat{\theta}\rangle$. In most systems encountered in practice, A_0 is not known and the entire estimate of A must be derived from the $\langle z \cos\hat{\theta}\rangle$ input.

The subsystem for estimating frequency and phase is also easily recognized as a second-order phase lock loop. The second integrator, whose input is $\hat{\theta}$, usually consists of a voltage controlled oscillator (VCO) which directly produces the $\sin\hat{\theta}$ and $\cos\hat{\theta}$ outputs. Except for the effect of the small feedback loop around the first integrator, the control function for the VCO is of the "proportional plus integral" type where the output $\langle z \sin\hat{\theta}\rangle$ of the phase detector is applied to both the first and second integrator inputs. The primary effect of the feedback around the first integrator is to limit its low-frequency gain to remain constant below α radians/second rather than continuing on to infinity at zero frequency. The mean frequency input ω_0 also serves to return the

frequency estimate to this value in the absence of other
information. Again this use of a known mean frequency is seldom
encountered in phase lock loop systems. However, it is a useful
concept whenever any a priori information about the center
frequency of the sinusoidal signal is available and provides
a means of "tuning" the detector prior to lock on, or of
prompting it to expected changes in average frequency such
as from doppler effects between transmitter and receiver.

STEADY STATE FORM OF FILTER

When a Kalman filter is used in problems where all defining matrices are stationary in time (or where short-term averages are stationary as in this problem), a steady state is eventually reached where P becomes constant and the gain coefficients in the state estimation part of the filter also become constant. Unless the behavior of the system before the steady state is reached is of particular importance, it is usually satisfactory to build a fixed parameter version of the state estimation filter using the steady state values of P and to avoid completely the necessity of implementing the differential equations for P. The steady state values of P also indicate the mean square errors associated with the fixed parameter estimation filter in the steady state and thus serve as a performance measure.

Finding the steady state values of P requires setting the equations (19) for $\dot{P}$ equal to zero and solving the resulting simultaneous nonlinear equations. Setting each of equations (19) to zero and rearranging slightly yields

$$2 \; P_{\theta\omega} = (A^2/2N) \; P_{\theta\theta}^2 \tag{23a}$$

$$\alpha^2 Q_\omega = 2\alpha P_{\omega\omega} + (A^2/2N) \; P_{\theta\omega}^2 \tag{23b}$$

$$\beta^2 Q_A = 2\beta P_{AA} + (1/2N) \; P_{AA}^2 \tag{23c}$$

$$P_{\omega\omega} = \alpha P_{\theta\omega} + (A^2/2N) P_{\theta\theta} P_{\theta\omega} \tag{23d}$$

Equation (23c) contains only P_{AA} and can be solved by itself as a quadratic in P_{AA} to yield

20

$$P_{AA} = -2\beta N \pm \sqrt{4 \beta^2 N^2 + 2\beta^2 N Q_A}$$

$$= 2\beta N[\sqrt{1 + (Q_A/2N)} - 1] \tag{24}$$

where the second root has been dropped as clearly extraneous. The remaining three equations can be solved simultaneously by using (23a) to substitute for $P_{\theta\omega}$ in (23b) and (23d) to give

$$\alpha^2 Q_\omega = 2\alpha P_{\omega\omega} + (A^2/2N)^3 P_{\theta\theta}^4/4 \tag{25a}$$

$$P_{\omega\omega} = \alpha(A^2/2N)P_{\theta\theta}^2/2 + (A^2/2N)^2 P_{\theta\theta}^3/2 \tag{25b}$$

Then substituting (25b) into (25a) gives

$$\alpha^2 Q_\omega = \alpha^2 \rho P_{\theta\theta}^2 + \alpha\rho^2 P_{\theta\theta}^3 + \rho^3 P_{\theta\theta}^4/4$$

$$= \rho P_{\theta\theta}^2 (\alpha + \rho P_{\theta\theta}/2)^2 \tag{26}$$

where

$$\rho = A^2/2N \tag{27}$$

is a signal to noise ratio (strictly signal power to noise power density, with the dimensions of Hertz) parameter. Multiplying by ρ and taking the square root on both sides converts equation (26) into a quadratic of the form

21

$$(\rho P_{\theta\theta})^2/2 + \alpha(\rho P_{\theta\theta}) \pm \alpha\sqrt{\rho Q_\omega} = 0 \tag{28}$$

for which the solutions are

$$P_{\theta\theta} = [-\alpha \pm \sqrt{\alpha^2 \pm 2\alpha\sqrt{\rho Q_\omega}}]/\rho$$

$$= (\alpha/\rho)[\sqrt{1 + 2\sqrt{\rho Q_\omega}/\alpha} - 1] \tag{29}$$

where only the single real positive solution has been retained.
Equation (25b) may now be used to find $P_{\omega\omega}$ as

$$P_{\omega\omega} = \alpha\rho(\alpha/\rho)^2[\sqrt{1 + 2\sqrt{\rho Q_\omega}/\alpha} - 1]^2/2$$

$$+ \rho^2(\alpha/\rho)^3[\sqrt{1 + 2\sqrt{\rho Q_\omega}/\alpha} - 1]^3/2$$

$$= (\alpha^3/2\rho)\sqrt{1 + 2\sqrt{\rho Q_\omega}/\alpha}\,[\sqrt{1 + 2\sqrt{\rho Q_\omega}/\alpha} - 1]^2 \tag{30}$$

and equation (23a) gives $P_{\theta\omega}$ as

$$P_{\theta\omega} = (\rho/2)(\alpha/\rho)^2[\sqrt{1 + 2\sqrt{\rho Q_\omega}/\alpha} - 1]^2$$

$$= (\alpha^2/2\rho)[\sqrt{1 + 2\sqrt{\rho Q_\omega}/\alpha} - 1]^2 \tag{31}$$

Equations (24), (29), (30), and (31) thus represent the steady
state solution for the uncertainty matrix.

22

These equations for the components of P do not provide
much physical insight when written in terms of Q_A and Q_ω.
They reduce to their simplest forms when some new variables
are introduced, representing a signal to noise ratio and
a modulation index for each of the modulation components.
The rms frequency deviation has been shown to be $D = \sqrt{\alpha Q_\omega/2}$,
and a frequency modulation index can be defined of the form

$$I_\omega = D/\alpha = \sqrt{Q_\omega/2\alpha} \tag{32}$$

This definition of the modulation index is consistent with the
usual (frequency deviation)/(modulating frequency) used
in describing FM systems since α represents a
bandwidth of the modulating process. The modulation index
can be loosely identified with an rms phase deviation, although
the phase is strictly a random walk process in this problem
with no finite variance.

Similarly an rms modulation percentage can be defined for
the amplitude modulation component by dividing the rms amplitude
fluctuation $\sqrt{\beta Q_A/2}$ by the mean amplitude of the signal. This
gives a parameter

$$I_A = \sqrt{\beta Q_A/2}/A \tag{33}$$

The signal to noise density ratio ρ can be
nondimensionalized for each modulation component by dividing
by the appropriate modulating process bandwidth. This gives

$$\rho_\omega = \rho/\alpha = A^2/2\alpha N \tag{34}$$

and

$$\rho_A = \rho/\beta = A^2/2\beta N \tag{35}$$

Physically these may be considered to be the non-dimensional
signal to noise ratios measured over the "natural" bandwidths
of the processes being estimated.

In terms of these new variables, the equations for the
P matrix components may be rewritten

$$P_{AA}/A^2 = [\sqrt{1 + 2\, I_A^2 \rho_A} - 1]/\rho_A \tag{36}$$

$$P_{\theta\theta} = [\sqrt{1 + 2\, I_\omega \sqrt{2\rho_\omega}} - 1]/\rho_\omega \tag{37}$$

$$P_{\omega\omega} = (\alpha^2/2)\sqrt{1 + 2\, I_\omega \sqrt{2\rho_\omega}}\,[\sqrt{1 + 2\, I_\omega \sqrt{2\rho_\omega}} - 1]^2/\rho_\omega \tag{38}$$

$$P_{\theta\omega} = (\alpha/2)[\sqrt{1 + 2 I_\omega \sqrt{2\rho_\omega}} - 1]^2/\rho_\omega \tag{39}$$

If the steady-state values of the P matrix as derived in
equations (36) through (39) are inserted as the appropriate

coefficients in the system shown in Figure 1, a steady-state
form of the Kalman filter can be derived as shown in Figure 2.
Figure 2a shows the amplitude estimation portion of the
system, and Figure 2b shows the frequency tracking portion.
The form shown for the amplitude estimator identifies it as a
simple low-pass exponential averager with a time constant τ_A
given by

$$\tau_A = 1 \Big/ \left[\beta \sqrt{1 + 2I_A^2 \rho_A} \right] \tag{40}$$

This averager has two inputs; one is the a priori average signal
level A_0 and the other is the measured quantity $<2z \cos\hat{\theta}>$.
The weight attached to the measured data input is

$$G_A = [\sqrt{1 + 2I_A^2 \rho_A} - 1] / \sqrt{1 + 2I_A^2 \rho_A} \tag{41}$$

while the weight attached to the a priori information is
$1 - G_A$. τ_A and G_A are plotted in Figure 3 as a function of
ρ_A for several values of the modulation index I_A. Notice
that if either the modulation index or the signal to noise ratio
is small, the time constant approaches the natural time
constant $1/\beta$ of the input fluctuations and most of the weight
is attached to the a priori estimate. In the limiting case the
output simply becomes equal to A_0 and the measured amplitudes
are ignored. For large values of I_A and ρ_A most of the input
weight goes to the measured amplitude $<2z \cos\hat{\theta}>$ and the time
constant of the averager becomes shorter. In the limit this time
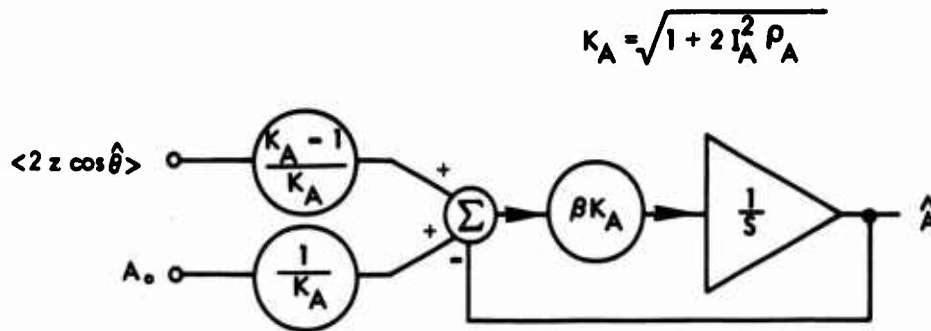constant approaches the asymptotic form
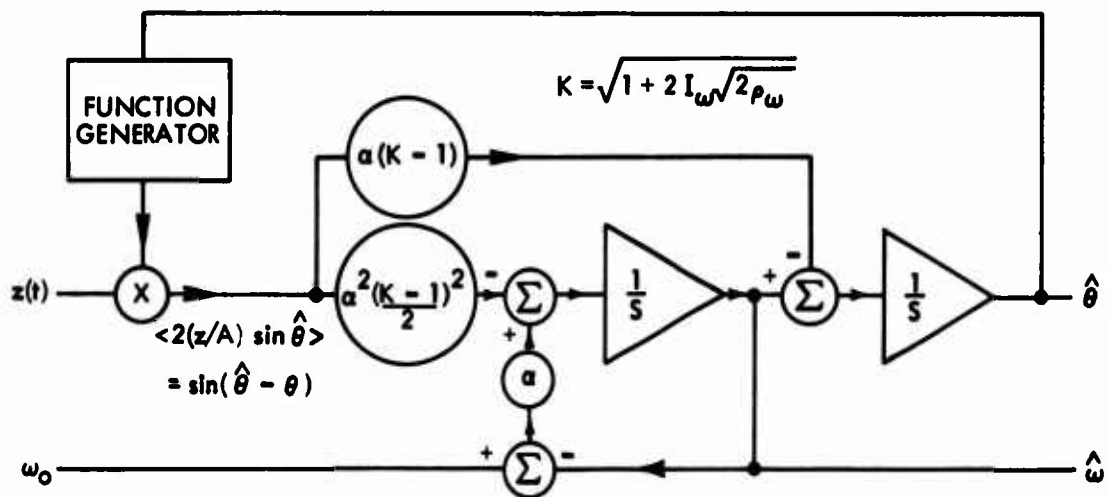
FIG. 2A AMPLITUDE ESTIMATION CIRCUIT

FIG. 2B PHASE AND FREQUENCY ESTIMATION CIRCUIT

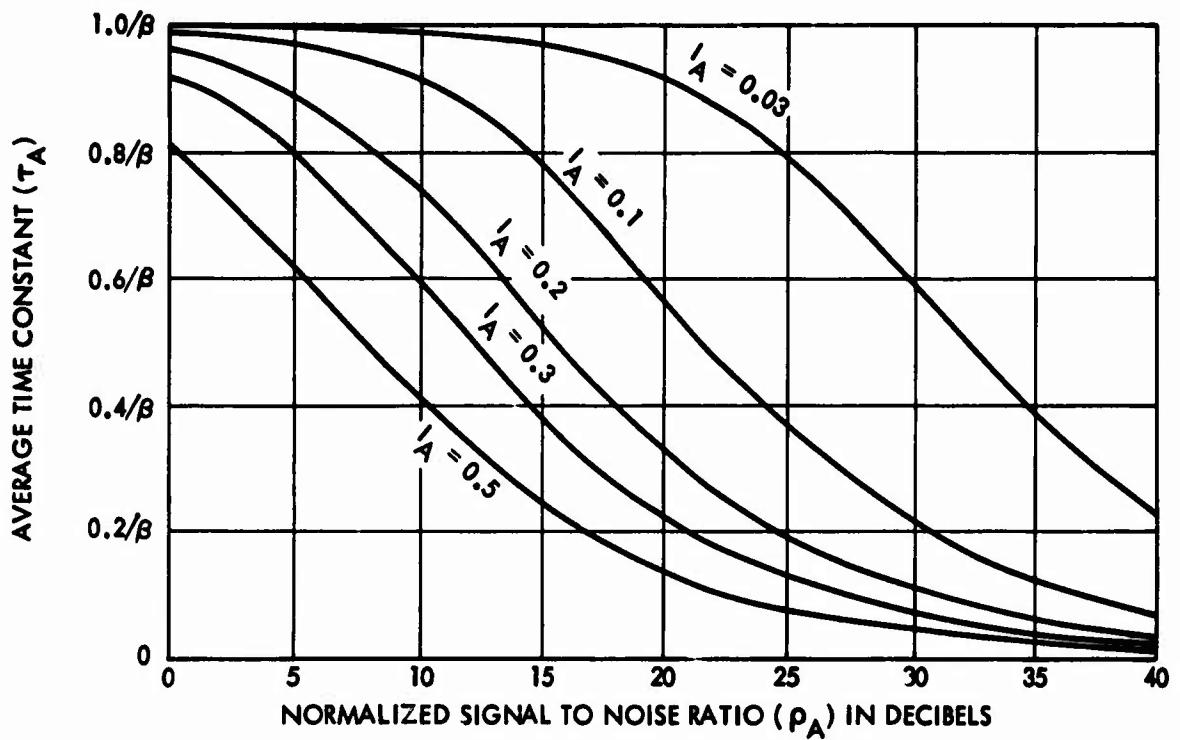FIG. 2 STEADY STATE FORM OF KALMAN FILTER
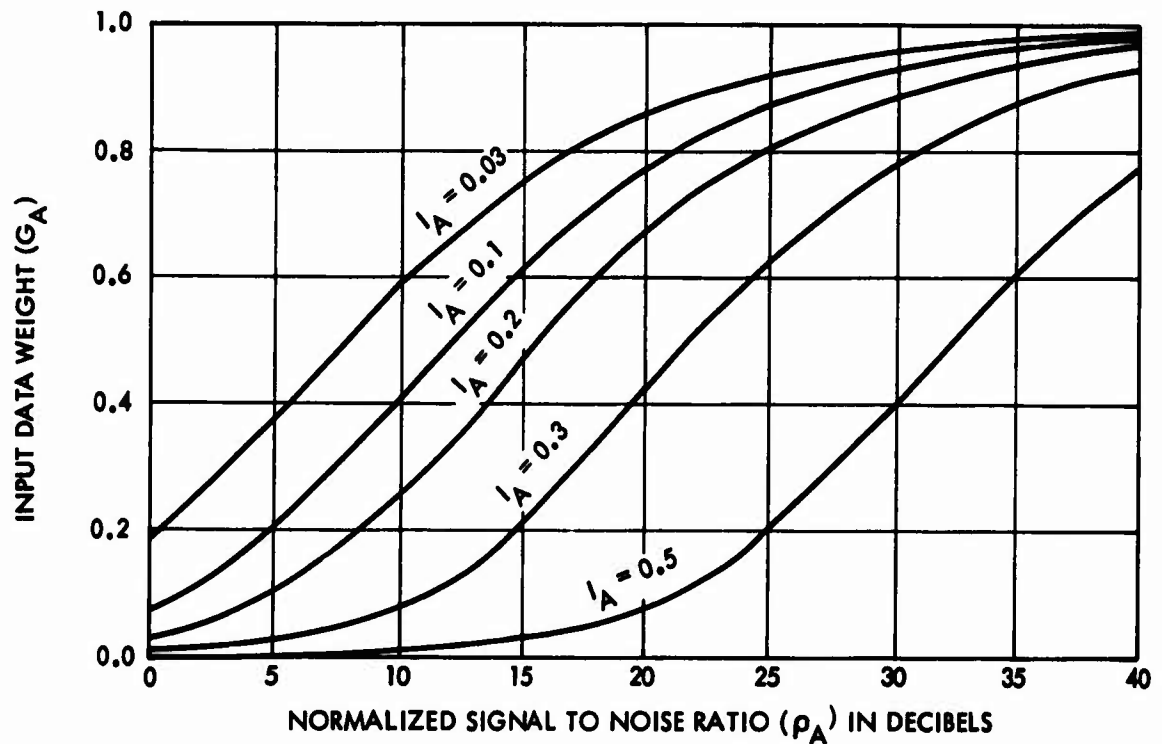
FIG. 3A VARIATION OF TIME CONSTANT

FIG. 3B VARIATION OF INPUT DATA WEIGHT

FIG. 3 VARIATION OF AVERAGER PARAMETERS WITH S/N AND MODULATION INDEX

$$\tau_A = 1/I_A\beta\sqrt{2\rho_A} \quad \text{for } I_A^2\rho_A \gg 1 \tag{42}$$

The frequency estimator shown in Figure 2b forms a phase lock loop circuit with a local feedback loop around the first integrator. This local feedback loop causes the first integrator to behave like a low-pass filter rather than a full integrator. A transfer function can be derived for the phase lock loop system by linearizing the output $\langle 2(z/A)\sin\hat{\theta}\rangle$ of the phase detector, such that this output is approximately $\hat{\theta} - \theta$ for small phase errors. The transfer function then becomes

$$T(s) = \frac{\hat{\theta}(s)}{\theta(s)} = \frac{2(K-1)\alpha s + (K^2-1)\alpha^2}{2s^2 + 2K\alpha s + (K^2-1)\alpha^2} \tag{43}$$

where

$$K = \sqrt{1 + 2I_\omega\sqrt{2\rho_\omega}} \tag{44}$$

The poles of this second order system are located at

$$s_{pole} = (\alpha/2)(-K \pm j\sqrt{K^2 - 2}) \tag{45}$$

which form a complex pair whenever $K > \sqrt{2}$. It is convenient to discuss the "bandwidth" of the phase lock loop system which may be defined as the radius to these poles. This bandwidth is

28

$$BW = \alpha \sqrt{(K^2-1)/2} = \alpha \sqrt{I_\omega} \sqrt{2\rho_\omega} \qquad (46)$$

and is plotted in Figure 4 for the region where the complex poles occur. It may be seen from equation (45) that for large K the real and imaginary parts of the pole location are equal (that is, the system is about 70% critically damped) and that the damping increases for smaller K until critical damping is reached for $K = \sqrt{2}$.

As the parameter K becomes smaller than $\sqrt{2}$ the inputs from the phase detector become progressively weaker, until for K very near unity all that is left is the local feedback loop around the ω integrator which causes the frequency estimate to settle at ω_0. This causes the system to degrade gracefully to the a priori estimate of the average frequency when there is insufficient information from the input to permit a better estimate.
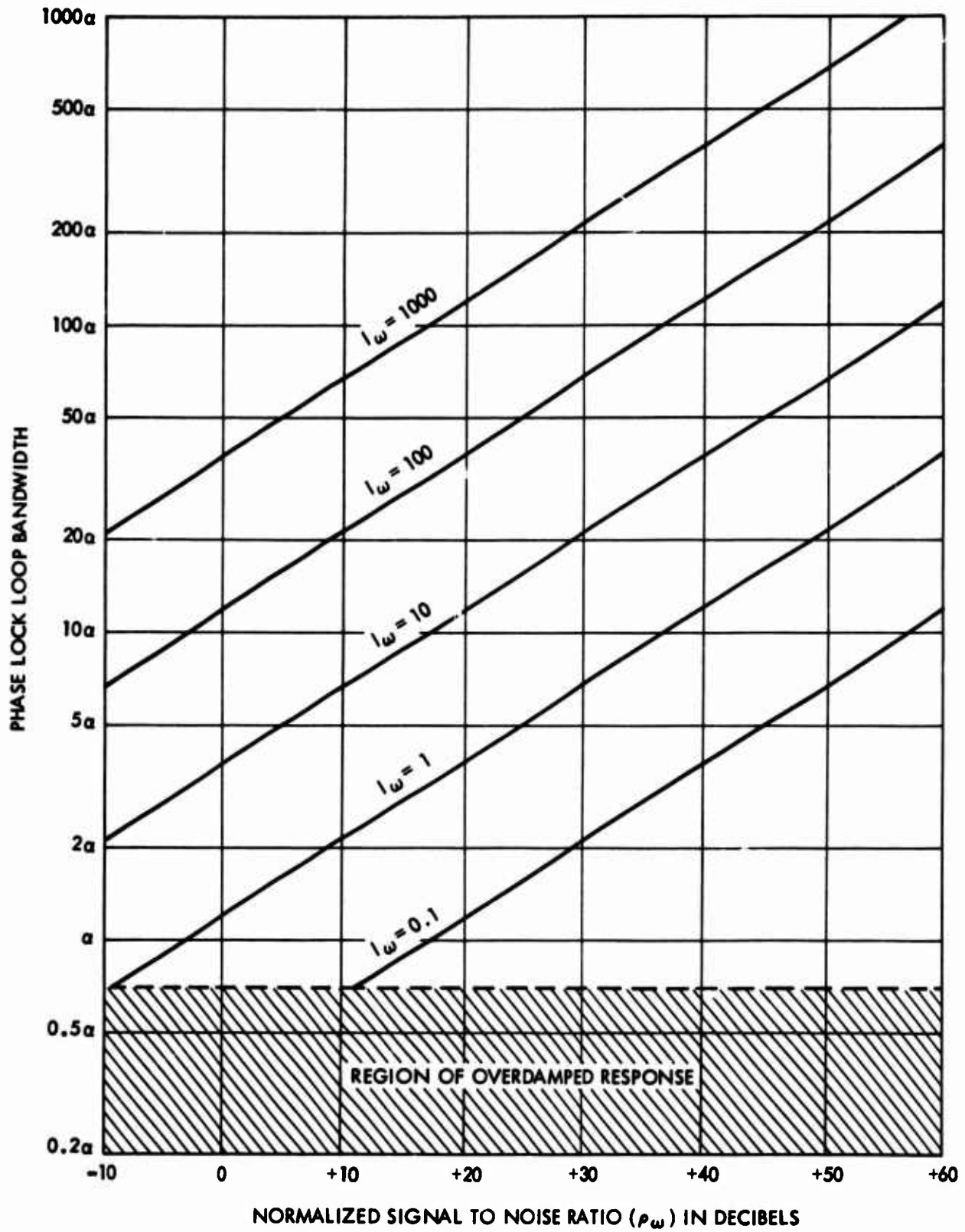
FIG. 4 BANDWIDTH OF PLL VS. S/N AND MODULATION INDEX

STEADY STATE PERFORMANCE

The steady state values of the P matrix components given in
equations (36), (37), and (38) also indicate the mean squared
errors in the amplitude, phase, and frequency estimates in the
steady state. Figures 5, 6, and 7 represent the rms errors expected
in each of the estimated quantities, in terms of the appropriate
signal to noise ratio and modulation index. Figure 5 shows
the relative uncertainty σ_A/A of the amplitude estimate versus
ρ_A and I_A, where the signal to noise ratio ρ_A is expressed in
decibel form. Note that in the limit of low ρ_A the relative
RMS error is just equal to the modulation index I_A. This is
consistent with the previously noted fact that the estimate
in this case is just equal to A_o so that the entire modulation
shows up as an error output. As the signal to noise ratio
increases the filter begins to track the amplitude fluctuations
so that the rms error begins to decrease. For a sufficiently
large signal to noise ratio, the amplitude estimation error
asymptotically approaches the form

$$\sigma_A/A = \sqrt{P_{AA}/A^2} \approx (I_A)^{1/2}(2/\rho_A)^{1/4} \quad \text{for } \rho_A \gg 1/I_A^2 \quad (47)$$

This somewhat strange dependence on I_A and ρ_A is due to the fact
that the time constant of the estimation filter is also a function
of these two parameters. Since the ratio σ_A/A has the same
significance at the output of the amplitude estimator as the
modulation index I_A at the input, an improvement factor equal to
the reduction in uncertainty of the signal amplitude may be defined
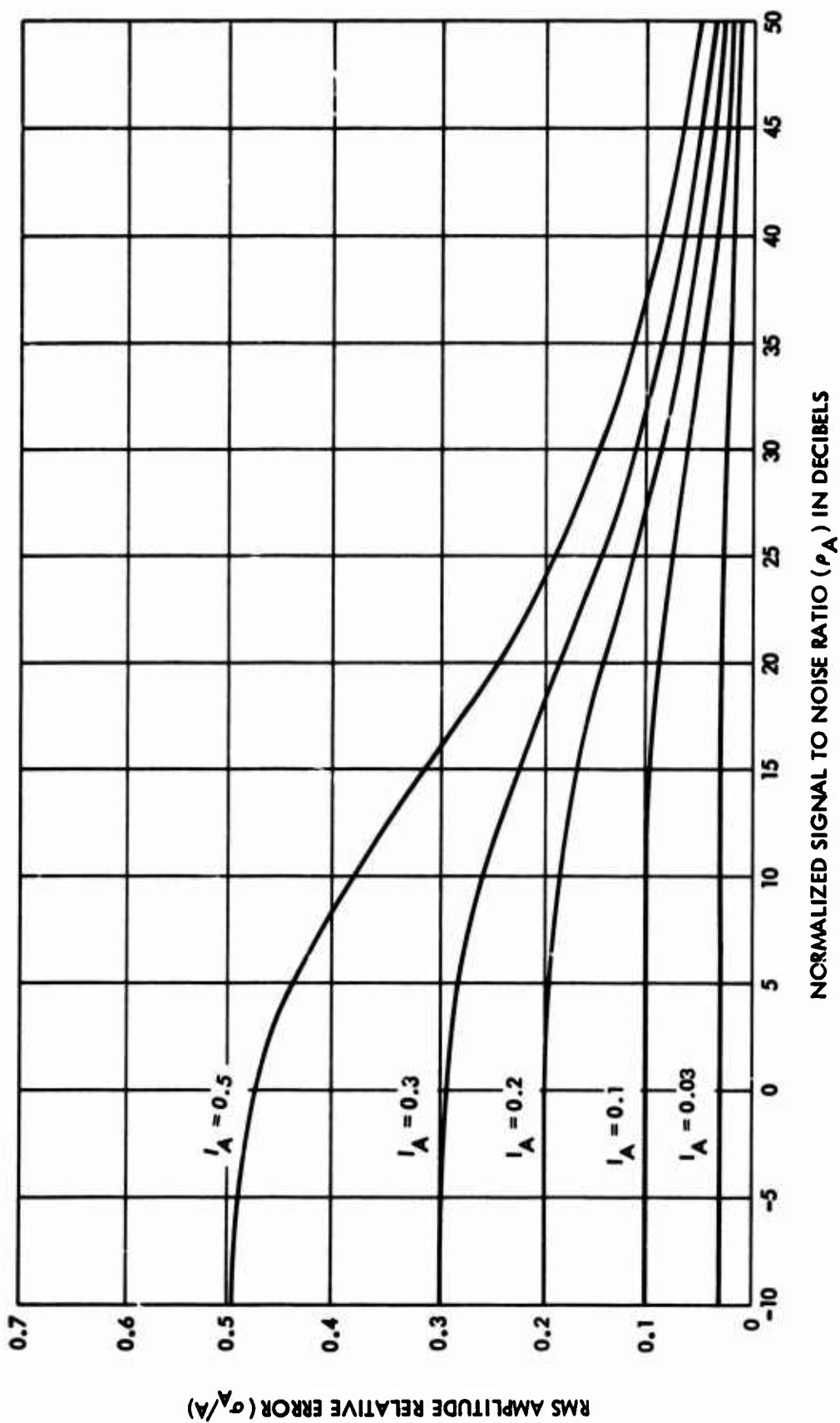and written in the form

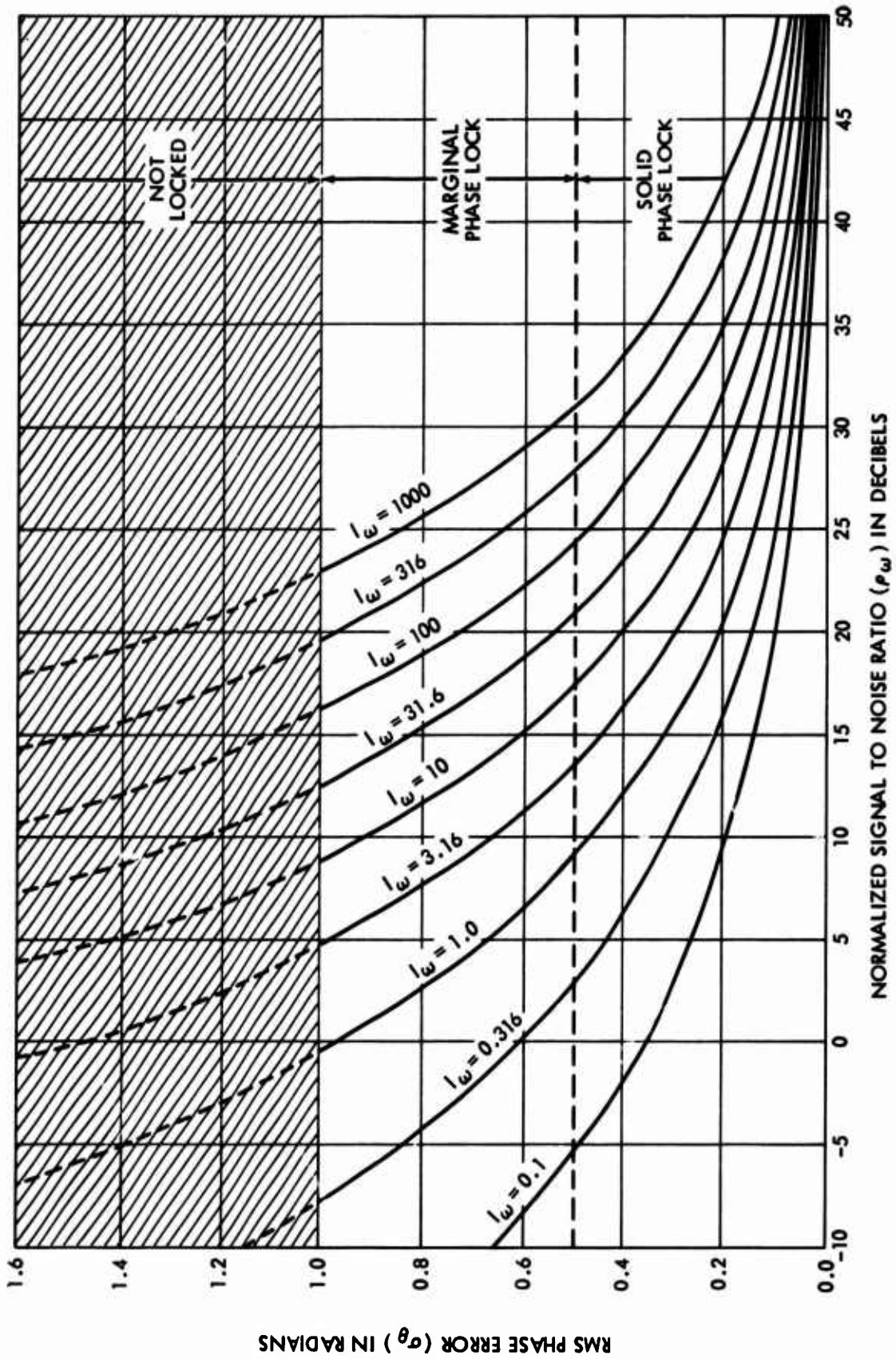FIG. 5 AMPLITUDE ESTIMATION ERROR VS. S/N AND MODULATION INDEX

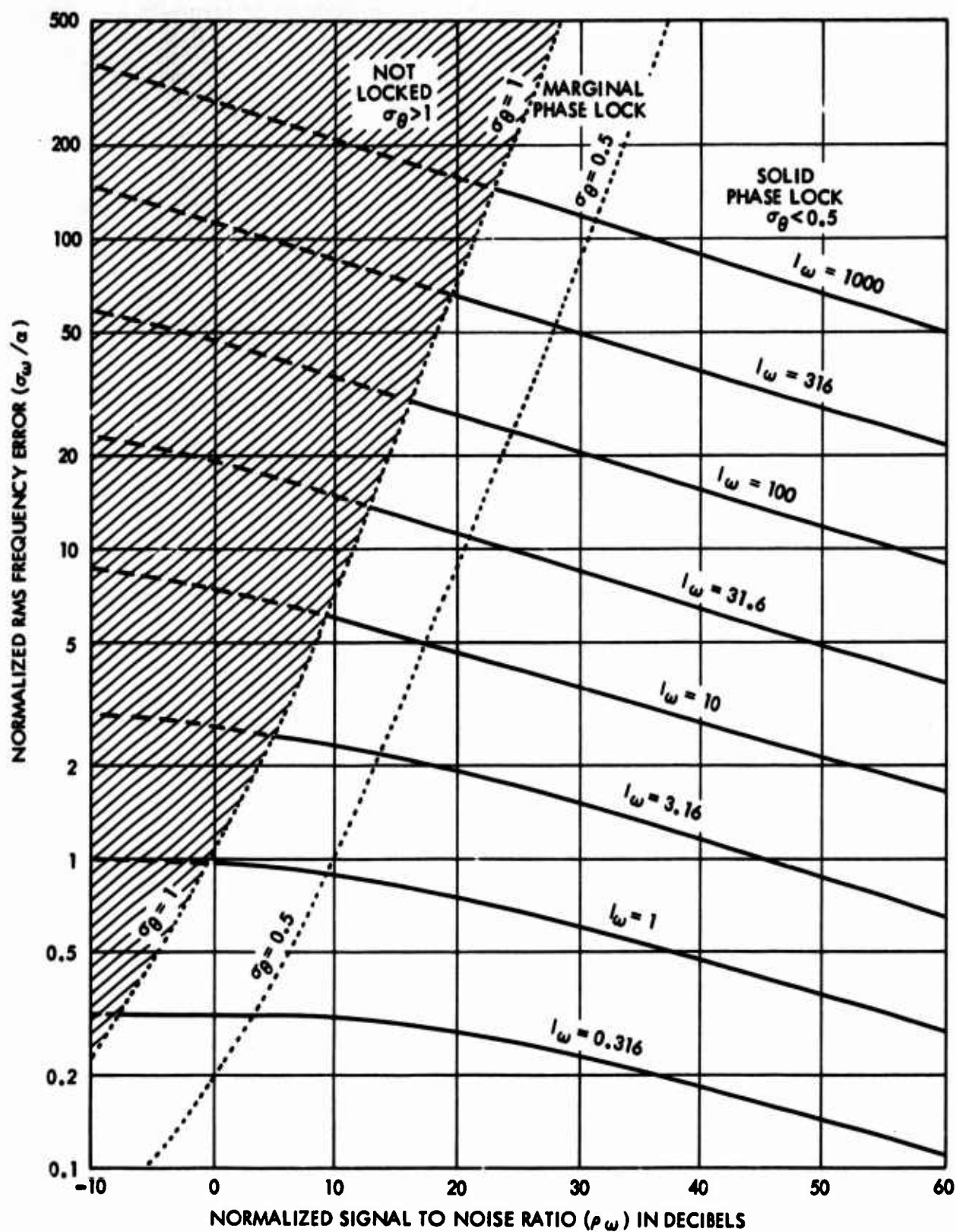FIG. 6 PHASE TRACKING ERROR VS. S/N AND MODULATION INDEX

FIG. 7 FREQUENCY TRACKING ERROR VS. S/N AND MODULATION INDEX

$$(\sigma_A/A)/I_A \simeq (2/I_A{}^2\rho_A)^{1/4} = (2\beta/I_A{}^2\rho)^{1/4}$$

$$\text{for } I_A{}^2\rho_A >> 1 \tag{48}$$

This factor represents the amount by which the rms uncertainty in the amplitude may be reduced by using the Kalman filter estimator, relative to just using the a priori estimate A_o.

Figure 6 shows the rms phase error of the phase and frequency estimation part of the system. The signal to noise ratio ρ_ω is expressed in decibels and the modulation index I_ω appears as a parameter. Note that the rms phase error is not bounded at low signal to noise ratios but continues to rise indefinitely. This is also a reasonable result, since the system shown in Figure 2b really gives up trying to track phase at low signal to noise ratios and the phase is a random walk process with infinite steady state variance. Once the rms phase error exceeds some upper limit it is inappropriate to discuss the estimation system as a phase lock loop, since the estimator is not truly "locked" to the input signal and the linearity assumption used in deriving equation (43) ceases to hold. If we arbitrarily choose an rms phase error of one radian as the upper limit for which we will consider the phase lock loop "locked", then a relationship of the form

$$I_\omega < (\rho_\omega/2)^{3/2} + (\rho_\omega/2)^{1/2} \tag{49}$$

can be derived as a criterion for lockon. For combinations
of modulation index and signal to noise ratio violating this
requirement the rms phase error exceeds one radian and the
main function of the phase detector output is to "pull" the
frequency estimate toward the true signal frequency without
actually following it. For a suffic'ently large signal
to noise ratio the rms phase error approaches an asymptotic form

$$\sigma_\theta = \sqrt{P_{\theta\theta}} \approx (I_\omega)^{1/4}(2/\rho_\omega)^{3/8} \quad \text{for } \rho_\omega \gg 1/8I_\omega^2 \quad (50)$$

The rms error in the frequency estimate is shown in
Figure 7 as a function of the signal to noise ratio ρ_ω
(expressed in decibels) and the modulation index I_ω. The rms
frequency error plotted along the vertical axis of Figure 7
is normalized by dividing by the "natural" fluctuation
frequency α to make it dimensionless. This normalized rms frequency
estimation error can be seen to approach I_ω as a limit for small
signal to noise ratios ρ_ω. Since I_ω is defined as the rms
frequency deviation D divided by α, this means that in the limit
of low signal to noise ratio the rms frequency error is just
equal to the rms deviation of the input frequency from the mean
ω_0. Again this result is physically reasonable, since we have seen
that the phase lock loop degenerates into a simple circuit whose
output is simply ω_0 for this case. In the limit of high signal to
noise ratios, the rms frequency estimation error approaches the
asymptotic form

$$\sigma_\omega = \sqrt{P_{\omega\omega}} \approx (2)^{1/2}(\alpha)^1(I_\omega)^{3/4}(2/\rho_\omega)^{1/8} \quad \text{for } \rho_\omega >> 1/8I_\omega^2 \tag{51}$$

It is instructive to re-introduce the rms input frequency deviation D in equation (51) so that the rms error can be expressed in terms of an improvement factor

$$\sigma_\omega/D \approx (32/I_\omega^2\rho_\omega)^{1/8} = (32\alpha^3/D^2\rho)^{1/8} \quad \text{for } 8\rho_\omega I_\omega^2 >> 1 \tag{52}$$

where the improvement factor is the ratio of the rms estimation error to the input frequency deviation. The improvement factor thus represents the reduction in signal frequency uncertainty possible by using the Kalman filter tracker to estimate the signal frequency, as compared to using only the a priori estimate ω_o for the signal frequency and accepting the rms deviation D as the uncertainty.

It is interesting that the improvement factor only goes as the eighth root of the signal to noise ratio, so that increased signal to noise ratio improves the tracking accuracy only very slowly. This result is partly due to the model used here for frequency fluctuations, where the random modulation process has significant power at high frequencies. Thus even with the large tracking bandwidths occurring at high signal to noise ratios, frequency fluctuations of still higher frequency occur and are not tracked. A signal model assuming smoother behavior of the frequency fluctuations would not necessarily give the same result.

The shaded area of Figure 7 represents the region where the rms phase error exceeds one radian and the loop is not locked in phase. Since the linearized Kalman filter derivation assumed small phase errors, this portion of the result violates these assumptions and is not necessarily correct. Some tendency to track the input frequency fluctuations is to be expected due to the "frequency pulling" effect in an unlocked phase lock loop, but it is not likely to be as good as that predicted by a linearized analysis. Thus the actual performance can be expected to follow the solid curves in the phase locked region but to degrade rapidly toward the no-signal performance of $\sigma_\omega/\alpha = I_\omega$ in going through the marginal phase lock region toward the unlocked region.

RECEIVER LOCKON THRESHOLD

As discussed in the previous section, for sufficiently low signal to noise ratios the rms phase errors are so large that the linearity assumptions used in deriving the receiver are violated. Since the receiver operation clearly relies on a phase lockon, it will fail to operate properly once the phase errors become excessively large. This is evidenced by failure to track the signal frequency and occasional "cycle slips" between the input signal and the receiver oscillator.

The choice of the maximum permissible phase error is somewhat arbitrary since tracking failures simply occur with increasing frequency as the rms phase error increases. However, a good case can be made for choosing σ_Θ = one radian as the lockon criterion. Recovery from phase errors as great as two radians is reasonably certain, but a phase error of π radians generally causes a cycle slip. Since errors of 2σ occur fairly frequently in a gaussian process while 3σ errors are comparatively rare, the performance of a phase lock receiver can be expected to degrade rapidly once σ_Θ exceeds one radian.

Equation (49) gives the relationship between I_ω and ρ_ω for an rms phase error below one radian. This can be converted to more physically meaningful terms by using equations (32) and (34) to give

$$D < \alpha \left[\left(\frac{\rho}{2\alpha} \right)^{3/2} + \left(\frac{\rho}{2\alpha} \right)^{1/2} \right] \tag{53}$$

In practical applications it is useful to express the signal to noise ratio (SNR) as the signal power divided by the noise power per Hertz of a one-sided (positive frequencies only) spectrum. Since ρ was defined in terms of the double-sided noise density N, the one-sided noise power per Hertz is 2N while the signal power is $A^2/2$, so we have

$$SNR = \frac{A^2/2}{2N} = \frac{\rho}{2} \tag{54}$$

Consequently for lockon with $\sigma_\theta <$ one radian we have

$$D < \alpha \left[(SNR/\alpha)^{3/2} + (SNR/\alpha)^{1/2} \right] = (\alpha \cdot SNR)^{1/2}(1 + SNR/\alpha) \tag{55}$$

This result is plotted in Figure 8, showing the minimum required SNR for tracking as a function of the rms frequency deviation D (expressed in Hertz rather than radians/second) for five values of modulation bandwidth α. For a given value of α, all combinations of SNR and D above the curve (higher SNR or lower D) result in acceptable phase lock while those below the curve do not.

It is interesting that for small frequency deviations lock-on occurs more readily for large values of α (rapid frequency fluctuations) than for small α, while for large frequency deviation signals with small values of α (long slow fluctuations) can be tracked at a lower SNR than can those with large α. There is good physical justification for this. If the modulation index
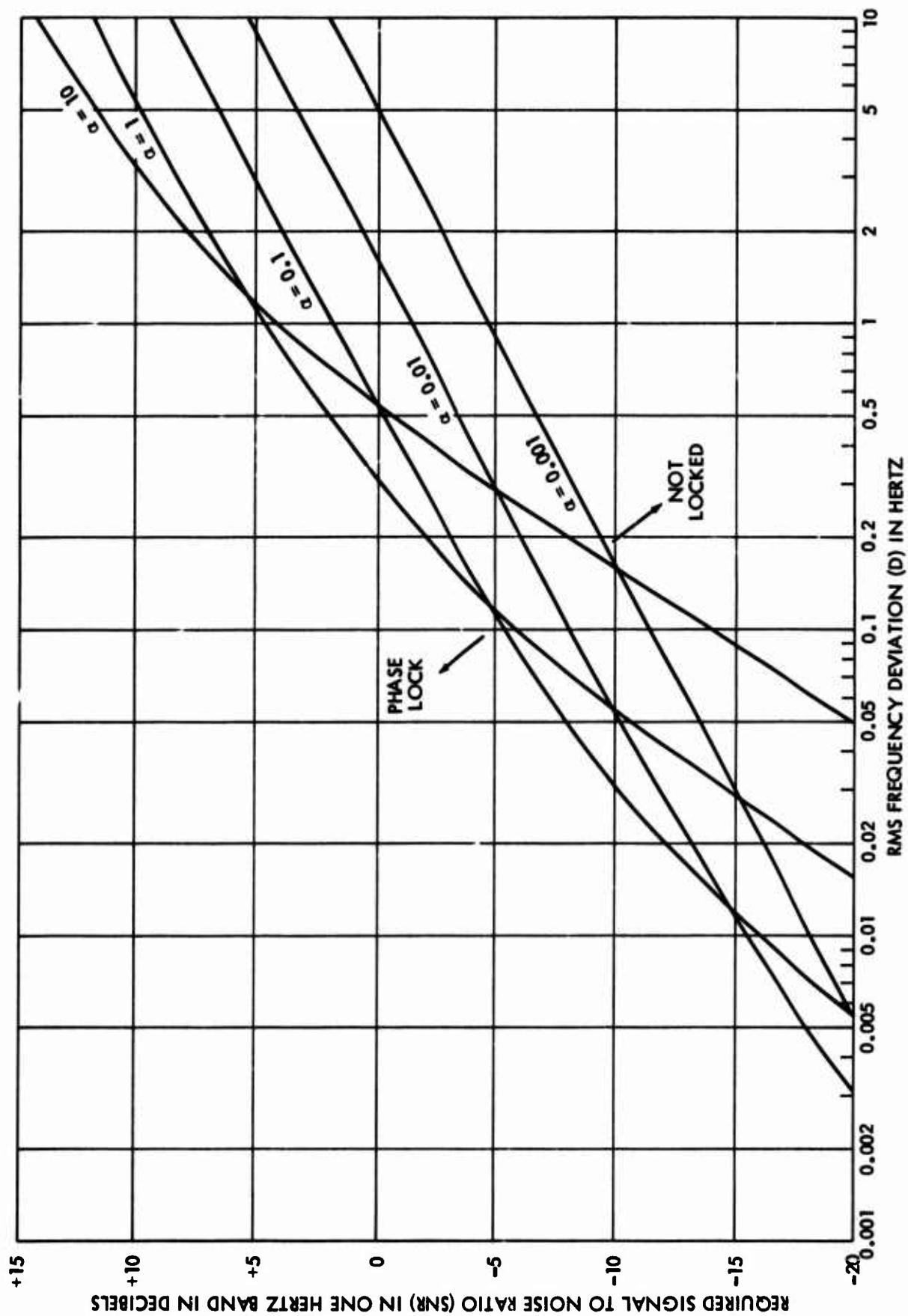
FIG. 8 REQUIRED SNR FOR PHASE LOCK VS. MODULATION BANDWIDTH AND DEVIATION

I_ω is small compared to 1, the system does not have to track the short term phase fluctuations and performs best when these fluctuations are rapid and soon average to zero. For I_ω greater than 1, the system must follow the phase changes to remain locked. It is then an advantage if they occur slowly.

AN EXAMPLE

Since the natural variables required to describe the frequency estimator performance are somewhat different from those commonly used in signal processing discussion, an example of the application of these results is appropriate. Suppose the sine wave to be tracked is known to have an rms frequency deviation of five Hertz about its center frequency with a "typical" relaxation time of one minute for the fluctuations. Suppose rms amplitude fluctuations of 30% are experienced about the average received amplitude and that these fluctuations are also correlated over about one minute. Then

$$\alpha = \beta = 1/60 \text{ radian/second}$$
$$I_A = 0.3$$
$$D = 5 \text{ Hertz} = 10\pi \text{ radians/second}$$
$$I_\omega = D/\alpha = 600\pi \simeq 1900$$

Notice that I_ω tends to be a rather large number whenever frequency fluctuations of this magnitude occur over periods of seconds or minutes. In fact, anytime one is inclined to describe the signal behavior as a sine wave of varying frequency rather than as a random narrow band process, this pretty much implies a value of I_ω much larger than unity. The amplitude modulation index I_A on the other hand is essentially constrained to be less than unity.

It is usual in signal processing problems of this class to represent the signal to noise ratio (SNR) as the total signal power

($A^2/2$) divided by the noise power in a one Hertz band, considering only positive real frequencies. Since the noise power density N in this treatment was defined for a two-sided spectrum, the noise in a one Hertz band of a one-sided spectrum is 2N. Thus the SNR is equal to $A^2/4N$ or $\rho/2$. The nondimensionalized signal to noise ratios ρ_A and ρ_ω are thus

$$\rho_A = \rho/\beta = 2(SNR)/\beta = 2(SNR)/(1/60) = 120(SNR)$$

$$\rho_\omega = \rho/\alpha = 120(SNR)$$

With these quantities thus defined we can use equations derived in the previous sections to determine system parameters and performance as a function of the SNR. In particular equation (40) gives the time constant of the amplitude estimator as

$$\tau_A = 1/\beta\sqrt{1 + 2I_A^2\rho_A}$$
$$= 60/\sqrt{1 + 2(0.3)^2(120)(SNR)}$$
$$= 60/\sqrt{1 + 21.6(SNR)} \text{ seconds} \tag{56}$$

equation (41) gives the weight for the measured amplitude information as

$$G_A = [\sqrt{1 + 2I_A^2\rho_A} - 1]/\sqrt{1 + 2I_A^2\rho_A}$$
$$= [\sqrt{1 + 21.6(SNR)} - 1]/\sqrt{1 + 21.6(SNR)} \tag{57}$$

and from equation (36) we obtain the rms relative error in
the estimated signal amplitude as

$$\sigma_A/A = \sqrt{P_{AA}/A^2} = \{[\sqrt{1 + 2I_A^2\rho_A} - 1]/\rho_A\}^{1/2}$$

$$= \{[\sqrt{1 + 21.6(SNR)} - 1]/120(SNR)\}^{1/2} \qquad (58)$$

These three results are plotted in Figure 9 as a function
of the SNR, expressed in decibels. Note that a SNR value
of about -13 db (that is SNR = 1/21.6) is the point
at which the amplitude estimator begins to work. Below
this point the input data is nearly ignored and the estimation error
is nearly equal to the input fluctuation amplitude. For higher
values of SNR the input measurements are more heavily weighted and
the time constant becomes shorter to track the fluctuations
more precisely.

Similarly for the frequency and phase estimation part of
the system, equation (46) gives the bandwidth of the phase lock
loop as

$$BW = \alpha\sqrt{I_\omega\sqrt{2\rho_\omega}} = (1/60)\sqrt{(1900)\sqrt{2(120)(SNR)}}$$

$$= 2.87(SNR)^{1/4} \text{ radians/second,} \qquad (59)$$

the rms phase error can be determined from equation (37) as
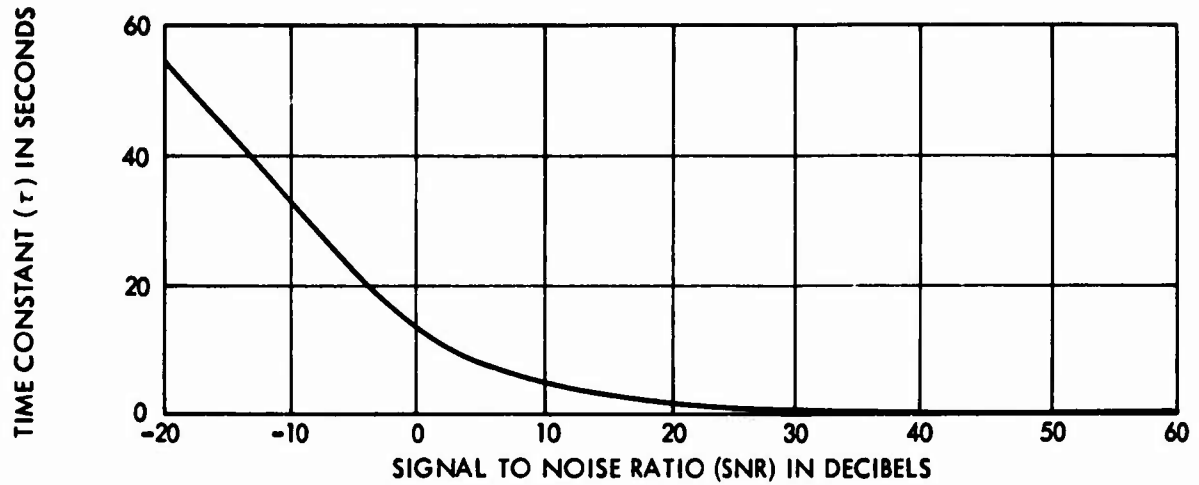
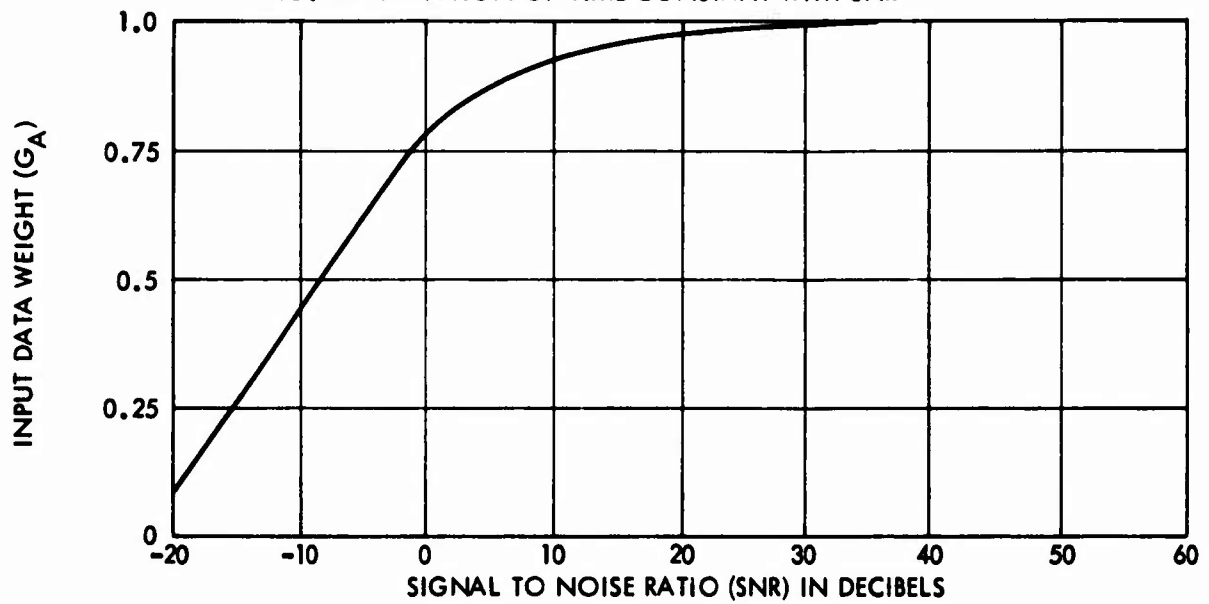FIG. 9A VARIATION OF TIME CONSTANT WITH SNR

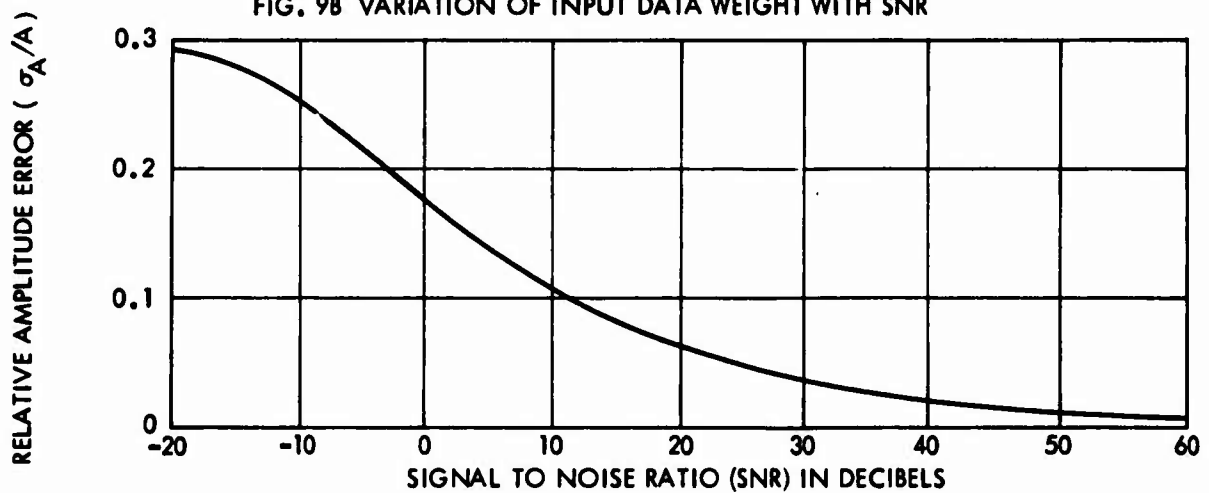FIG. 9B VARIATION OF INPUT DATA WEIGHT WITH SNR

FIG. 9C VARIATION OF RMS AMPLITUDE ERROR WITH SNR

FIG. 9 PARAMETERS AND PERFORMANCE OF EXAMPLE AMPLITUDE ESTIMATOR VS. SNR

$$\sigma_\theta = \sqrt{P_{\theta\theta}} = \{[\sqrt{1 + 2I_\omega\sqrt{2\rho_\omega}} - 1]/\rho_\omega\}^{1/2}$$

$$= \{[\sqrt{1 + 2(1900)\sqrt{2(120)(SNR)}} - 1]/(120)(SNR)\}^{1/2}$$

$$= \{[\sqrt{1 + (5.9 \ 10^4)(SNR)^{1/2}} - 1]/(120)(SNR)\}^{1/2} \qquad (60)$$

and the rms frequency error is obtained from equation (38) as

$$\sigma_\omega = \sqrt{P_{\omega\omega}} = \alpha[\sqrt{1 + 2I_\omega\sqrt{2\rho_\omega}} - 1] \cdot (1 + 2I_\omega\sqrt{2\rho_\omega})^{1/4}/(2\rho_\omega)^{1/2}$$

$$= (1/60)[\sqrt{1 + (5.9 \ 10^4)(SNR)^{1/2}} - 1] \cdot$$

$$(1 + 5.9 \ 10^4)(SNR)^{1/2})^{1/4}/(240(SNR))^{1/2}$$

$$= \frac{[\sqrt{1 + (5.9 \ 10^4)(SNR)^{1/2}} - 1] \cdot (1 + (5.9 \ 10^4)(SNR)^{1/2})^{1/4}}{930(SNR)^{1/2}} \qquad (61)$$

These three functions are plotted in Figure 10 as a function of the SNR, expressed in decibels. Note that the requirement

$$= 120(SNR) > 1/8I^2 = 1/8(1900)^2$$

is satisfied at an SNR of $2.88 \ 10^{-10}$ or about -96 db, so for all reasonable values of SNR the asymptotic forms in equations (50) and (51) hold to give

$$\sigma_\theta \approx (I_\omega)^{1/4}(2/\rho_\omega)^{3/8} = (1900)^{1/4}(60(SNR))^{-3/8}$$

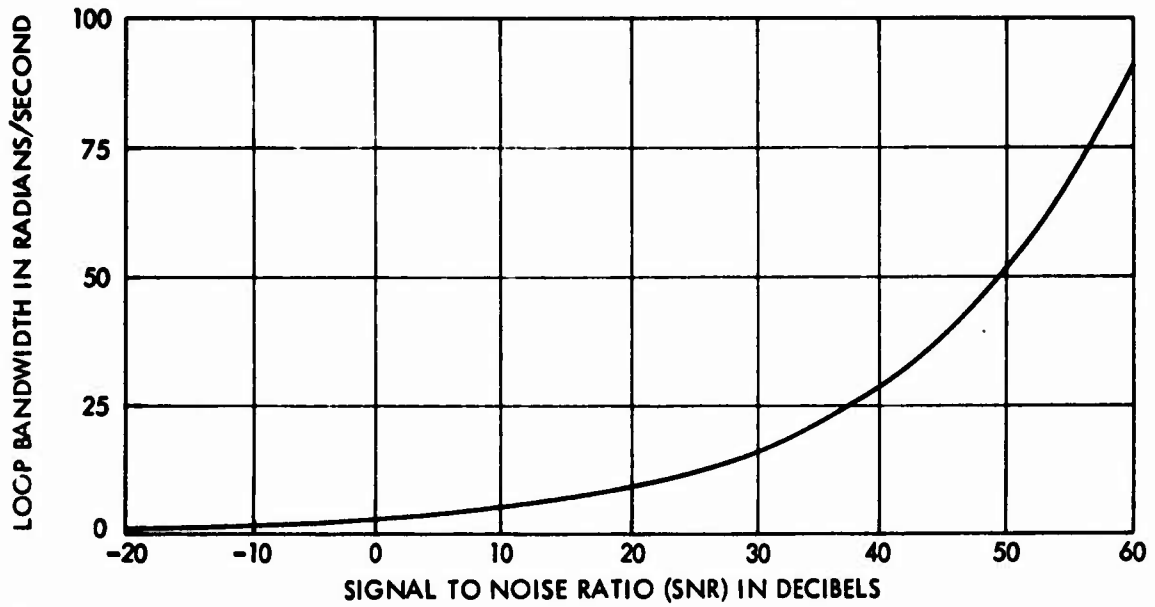$$= 1.42/(SNR)^{3/8} \text{ radians} \qquad (62)$$

and

47

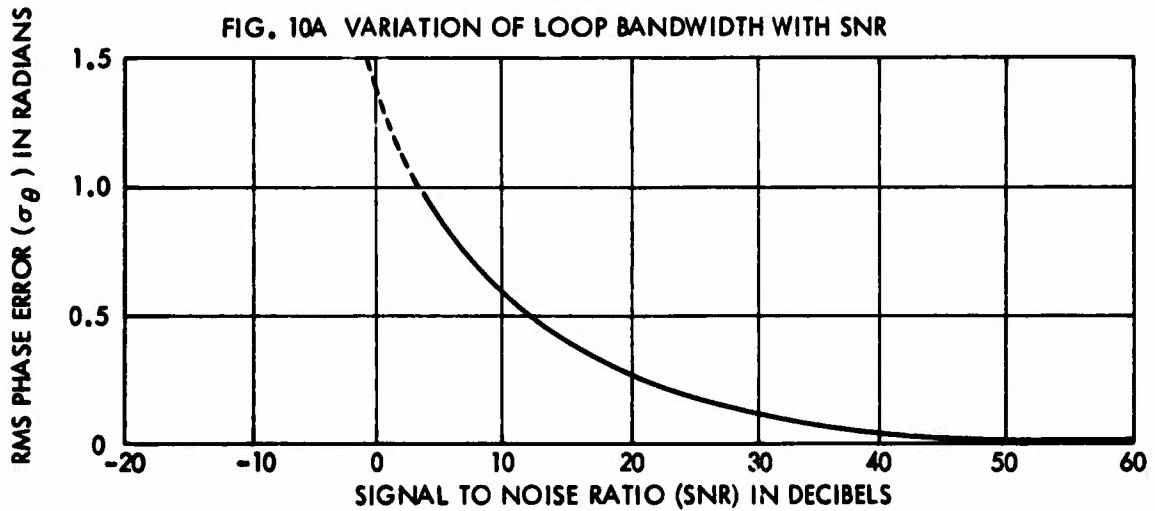FIG. 10A VARIATION OF LOOP BANDWIDTH WITH SNR

FIG. 10B VARIATION OF RMS PHASE ERROR WITH SNR

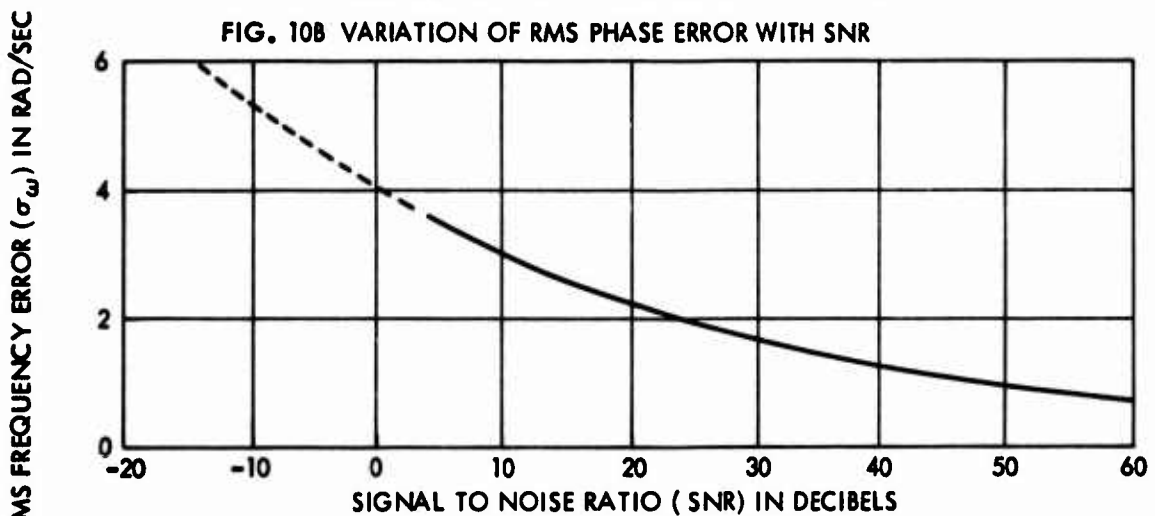FIG. 10C VARIATION OF RMS FREQUENCY ERROR WITH SNR

FIG. 10 PARAMETERS AND PERFORMANCE OF EXAMPLE FREQUENCY ESTIMATOR VS. SNR

$$\sigma_\omega = \alpha(2)^{1/2}(I_\omega)^{3/4}(2/\rho_\omega)^{1/8}$$

$$= (1/60)(1.414)(1900)^{3/4}(60(SNR))^{-1/8}$$

$$= 4.06(SNR)^{1/8} \text{ radians/second} \tag{63}$$

The curve for σ_θ shows the rms phase error becoming less than
one radian for SNR values above 2.52 or about +4 db. This
can be considered the threshold at which phase lock occurs.
Below this SNR, the expressions for frequency and phase
errors are probably meaningless and are shown dashed in
Figure 10. This minimum SNR point could have been determined
directly from Figure 8 using D = 5 Hertz and α = .016. Note
that because of the very weak dependence of σ_ω on the SNR,
the rms frequency estimation error is between about 1 and
4 radians per second from this lock on SNR to about +50 db.
Again it should be remembered that this weak dependence on
SNR is due in part to the model assumed for the signal
fluctuations. It should also be remembered that the curves
in Figures 9 and 10 are not for a fixed parameter system
operating over a range of SNR values. They represent a system
that for each SNR has the optimum parameters for that SNR
value, due to the amplitude dependence of the gain coefficients
shown in Figure 2.

SUMMARY AND CONCLUSIONS

This paper has addressed the problem of finding the optimum frequency estimator for a sinusoid whose frequency is fluctuating randomly in time and whose amplitude is fluctuating, in the presence of additive random noise. This is essentially the problem of demodulation of a frequency modulated signal through a noisy, fading transmission channel. The desired signal properties were modeled using continuous state variable methods and the optimum estimator derived using a linearized form of continuous Kalman filter theory. The problem is sufficiently small to permit expansion of the matrix equations for the estimator, to allow determination of the form of the optimum estimator. Once the unimportant high-frequency terms are removed, the resultant system is shown to resemble a phase lock loop for frequency tracking with a coherent detector and low-pass filter for amplitude estimation. A more complex model, using Rayleigh fading in the transmission channel, is covered in the Appendix. It is shown that the optimum estimator is again a phase lock loop, but both phases of the signal amplitude must be detected and used to phase-correct the phase lock loop reference signal. In both cases it is satisfying to find that the mathematically optimum processor is essentially the same as one would have been tempted to design on an ad hoc basis.

It is also possible to solve the differential equations
for the uncertainty matrix for the steady state case, in terms
of the input signal and noise parameters. This permits one to
determine the optimum parameters, as well as the form, of the
processor and to determine the steady state estimation errors.
Thus for the steady state case we have fully determined the
design of the optimum processor and its performance, as a
function of the characteristics of its inputs. These can either
be used directly in design of an estimator or can serve as a
reference against which an ad hoc processor can be compared.

The optimum parameters of the processor depend upon the
modulation characteristics of the signal (deviation and bandwidth
of the frequency modulation, percentage modulation and bandwidth
of the amplitude modulation) and on the signal to noise ratio.
Thus any particular set of parameters chosen represents a point
design for a given set of signal characteristics. However, since
the signal characteristics can all be measured by the processor,
the information derived here can also be used to design an
adaptive estimator which measures the characteristics of the
received signal and adjusts its parameters accordingly.

APPENDIX

EFFECT OF RAYLEIGH FADING SIGNAL MODEL

The signal model used in the main body of this report
assumed a phase-continuous signal of varying frequency, with
random modulation in amplitude. It was shown that the optimum
frequency estimator for this problem consisted of a phase lock
loop system. It can be argued that this result is sensitive
to the signal model, and that the phase lock loop is tracking
only the "carrier" portion of the amplitude modulated signal.
Thus the amplitude fluctuations have no effect on the resultant
processor. The sorts of physical phenomena for which this
signal model might be appropriate are those in which the signal
source itself fluctuates in amplitude, its coupling to the
transmission medium is highly time or aspect dependent as with
complex radiation patterns, or the transmission path itself
has a time dependent gain such as with moving absorptive cloud
masses.

In another broad class of problems, the amplitude
fluctuation is due to variations in the transmission path
due to multipath propagation. In both sky-wave high frequency
radio propagation and in underwater acoustic propagation, it
is common for the signal to arrive by several paths of
approximately equal strength but whose path lengths differ

by some number of wavelengths. As these signals interfere with
each other the resultant amplitude varies, but the phase also
varies due to the vector addition of signal arrivals of randomly
varying phase relationship. The most common analytical form
used for such signals assumes that each arrival path of the
signal consists of one random component of sine phase and one
of cosine phase. Then if many such signals are added, the
Central Limit Theorem indicates that the amplitudes of the
resultant sine and cosine components are independent zero-mean
Gaussian variables. From this model, the total amplitude may be
shown to follow a so-called Rayleigh distribution while the
resultant phase angle is a uniform random variable. This
Rayleigh model agrees well with experimental evidence and
includes such complications as rapid reversal of signal phase
at signal dropouts. Since the phase is not continuous and there
is no "carrier" to track, one would expect that this signal
model would have some effect on the form of the optimum receiver.

Following the same procedure as in the main body of the
report, the received signal state can be represented by the vector

$$
X] = \begin{bmatrix} \theta \\ \omega \\ A \\ B \end{bmatrix}
\tag{64}
$$

such that the measurable signal at any instant is

$$h(X) = h(\theta,\omega,A,B) = A\cos\theta + B\sin\theta \qquad (65)$$

A and B are thus the random amplitudes of the two components of the received signal and are zero-mean processes of equal variance. Again the frequency parameter ω does not directly appear in the measurable quantity.

There is another more serious problem with this formulation. Since h(X), and its evolution with time, can be represented by many combinations of A, B, and θ (for example adding π to the initial value of θ and reversing the signs of A and B), the true state of the system is not observable. This inherent coupling of the uncertainties of the amplitude and phase variables eventually causes analytical difficulties in the solution of this problem. However it is still possible to determine the form of the optimum receiver.

Again we need to represent the evolution of the state vector X by an equation of the form $\dot{X} = FX+Gw$, and the appropriate matrices are

$$F = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -\alpha & 0 & 0 \\ 0 & 0 & -\beta & 0 \\ 0 & 0 & 0 & -\beta \end{bmatrix} \qquad (66)$$

and

$$G = \begin{bmatrix} 0 & 0 & 0 \\ \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \beta \end{bmatrix} \tag{67}$$

The w vector is a three component random process with mean value and covariance given by

$$w_o = \begin{bmatrix} \omega_o \\ 0 \\ 0 \end{bmatrix} \tag{68}$$

and

$$Q = \begin{bmatrix} Q_\omega & 0 & 0 \\ 0 & Q_A & 0 \\ 0 & 0 & Q_A \end{bmatrix} \tag{69}$$

These definitions of F, G, and w again establish a system with a source center frequency of ω_o with random low-pass frequency modulation of "bandwidth" α. The rms deviation from the center frequency is again given by $\sqrt{\alpha Q_\omega/2}$. The complex amplitude modulation is a low-pass Rayleigh process of bandwidth β, and the mean power in the signal is determined by the value of Q_A. Note the mean amplitude and the rms fluctuations in amplitude are not separately specified as before, since they are inherently related by the Rayleigh distribution.

Additional matrices to be defined as inputs to the Kalman filter solution are the measurement noise matrix R, which is a 1x1 matrix equal to the additive noise power density N, the linearized measurement matrix H, which is

$$[H] = [\partial h/\partial\hat{\theta}, \; \partial h/\partial\hat{\omega}, \; \partial h/\partial\hat{A}, \; \partial h/\partial\hat{B}]$$

$$= [-\hat{A}\sin\hat{\theta}+\hat{B}\cos\hat{\theta}, \; 0, \; \cos\hat{\theta}, \; \sin\hat{\theta}] \tag{70}$$

and the estimation uncertainty matrix P, which has the form

$$[P] = \begin{bmatrix} P_{\theta\theta} & P_{\theta\omega} & P_{\theta A} & P_{\theta B} \\ P_{\omega\theta} & P_{\omega\omega} & P_{\omega A} & P_{\omega B} \\ P_{A\theta} & P_{A\omega} & P_{AA} & P_{AB} \\ P_{B\theta} & P_{B\omega} & P_{BA} & P_{BB} \end{bmatrix} \tag{71}$$

From these matrix definitions it is now possible in principle to derive the optimum Kalman filter estimator and to determine the performance from the steady state values of the P matrix. Again this is considerably harder than it sounds.

Following the same procedure as in the main body of this report, the first step is to expand the differential equation for P, so that

$$\dot{P} = F P + P F^T + G Q G^T - P H^T R^{-1} H P$$

$$= \begin{bmatrix} 2P_{\theta\omega} & P_{\omega\omega}-\alpha P_{\theta\omega} & P_{\omega A}-\beta P_{\theta A} & P_{\omega B}-\beta P_{\theta B} \\ P_{\omega\omega}-\alpha P_{\theta\omega} & \alpha^2 Q_\omega - 2\alpha P_{\omega\omega} & -\alpha P_{\omega A}-\beta P_{\omega A} & -\alpha P_{\omega B}-\beta P_{\omega B} \\ P_{\omega A}-\beta P_{\theta A} & -\alpha P_{\omega A}-\beta P_{\omega A} & \beta^2 Q_A - 2\beta P_{AA} & -2\beta P_{AB} \\ P_{\omega B}-\beta P_{\theta B} & -\alpha P_{\omega B}-\beta P_{\omega B} & -2\beta P_{AB} & \beta^2 Q_A - 2\beta P_{BB} \end{bmatrix}$$

$$-(1/N) \begin{bmatrix} P_{\theta\theta}(B\cos\theta-A\sin\theta) + P_{\theta A}\cos\theta + P_{\theta B}\sin\theta \\ P_{\theta\omega}(B\cos\theta-A\sin\theta) + P_{\omega A}\cos\theta + P_{\omega B}\sin\theta \\ P_{\theta A}(B\cos\theta-A\sin\theta) + P_{AA}\cos\theta + P_{AB}\sin\theta \\ P_{\theta B}(B\cos\theta-A\sin\theta) + P_{AB}\cos\theta + P_{BB}\sin\theta \end{bmatrix}$$

$$\begin{bmatrix} P_{\theta\theta}(B\cos\theta-A\sin\theta) + P_{\theta A}\cos\theta + P_{\theta B}\sin\theta \\ P_{\theta\omega}(B\cos\theta-A\sin\theta) + P_{\omega A}\cos\theta + P_{\omega B}\sin\theta \\ P_{\theta A}(B\cos\theta-A\sin\theta) + P_{AA}\cos\theta + P_{AB}\sin\theta \\ P_{\theta B}(B\cos\theta-A\sin\theta) + P_{AB}\cos\theta + P_{BB}\sin\theta \end{bmatrix}^T \qquad (72)$$

In the above, advantage has been taken of the fact that $P_{ij} = P_{ji}$ to simplify the equations slightly.

We next wish to take advantage of the fact that, once the indicated multiplication is performed in equation (72) above, many of the expressions will contain terms oscillating at $2\omega_o$.

Since these do not influence the long-term evolution of the P
matrix, they will be dropped. The resultant ten scalar equations
for the elements of P are:

$$\langle \dot{P}_{\theta\theta} \rangle = 2P_{\theta\omega} \quad -(1/2N)\left[(A^2+B^2)P_{\theta\theta}^2 + 2BP_{\theta\theta}P_{\theta A} - 2AP_{\theta\theta}P_{\theta B} + P_{\theta A}^2 + P_{\theta B}^2\right] \tag{73a}$$

$$\langle \dot{P}_{\omega\omega} \rangle = \alpha^2 Q_\omega - 2\alpha P_{\omega\omega} - (1/2N)\left[(A^2+B^2)P_{\theta\omega}^2 + 2BP_{\theta\omega}P_{\omega A} - 2AP_{\theta\omega}P_{\omega B} + P_{\omega A}^2 + P_{\omega B}^2\right] \tag{73b}$$

$$\langle \dot{P}_{AA} \rangle = \beta^2 Q_A - 2\beta P_{AA} - (1/2N)\left[(A^2+B^2)P_{\theta A}^2 + 2BP_{\theta A}P_{AA} - 2AP_{\theta A}P_{AB} + P_{AA}^2 + P_{AB}^2\right] \tag{73c}$$

$$\langle \dot{P}_{BB} \rangle = \beta^2 Q_A - 2\beta P_{BB} - (1/2N)\left[(A^2+B^2)P_{\theta B}^2 + 2BP_{\theta B}P_{AB} - 2AP_{\theta B}P_{BB} + P_{AB}^2 + P_{BB}^2\right] \tag{73d}$$

$$\langle \dot{P}_{\theta\omega} \rangle = P_{\omega\omega} - \alpha P_{\theta\omega} - (1/2N)\left[(A^2+B^2)P_{\theta\theta}P_{\theta\omega} + B(P_{\theta A}P_{\theta\omega} + P_{\theta\theta}P_{\omega A}) - A(P_{\theta B}P_{\theta\omega} + P_{\theta\theta}P_{\omega B}) + P_{\theta A}P_{\omega A} + P_{\theta B}P_{\omega B}\right] \tag{73e}$$

$$\langle \dot{P}_{\theta A} \rangle = P_{\omega A} - \beta P_{\theta A} - (1/2N)\left[(A^2+B^2)P_{\theta\theta}P_{\theta A} + B(P_{\theta\theta}P_{AA} + P_{\theta A}^2) - A(P_{\theta B}P_{\theta A} + P_{\theta\theta}P_{AB}) + P_{\theta A}P_{AA} + P_{\theta B}P_{AB}\right] \tag{73f}$$

$$\langle \dot{P}_{\theta B} \rangle = P_{\omega B} - \beta P_{\theta B} - (1/2N)\left[(A^2+B^2)P_{\theta\theta}P_{\theta B} + B(P_{\theta\theta}P_{AB} + P_{\theta A}P_{\theta B}) - A(P_{\theta\theta}P_{BB} + P_{\theta B}^2) + P_{\theta A}P_{\theta B} + P_{\theta B}P_{BB}\right] \tag{73g}$$

$$\langle \dot{P}_{\omega A} \rangle = -\alpha P_{\omega A} - \beta P_{\omega A} - (1/2N)\left[(A^2+B^2)P_{\theta\omega}P_{\theta A} + B(P_{\theta A}P_{\omega A} + P_{\theta\omega}P_{AA}) - A(P_{\theta A}P_{\omega B} + P_{\theta\omega}P_{AB}) + P_{\omega A}P_{AA} + P_{\omega B}P_{AB}\right] \tag{73h}$$

$$\langle \dot{P}_{\omega B} \rangle = -\alpha P_{\omega B} - \beta P_{\omega B} - (1/2N)\left[(A^2+B^2)P_{\theta\omega}P_{\theta B} + B(P_{\theta\omega}P_{AB}+P_{\omega A}P_{\theta B}) - A(P_{\omega B}P_{\theta B}+P_{\theta\omega}P_{BB}) + P_{\omega A}P_{AB} + P_{\omega B}P_{BB}\right] \quad (73I)$$

$$\langle \dot{P}_{AB} \rangle = -2\beta P_{AB} - (1/2N)\left[(A^2+B^2)P_{\theta A}P_{\theta B} + B(P_{\theta A}P_{AB}+P_{AA}P_{\theta B}) - A(P_{AB}P_{\theta B}+P_{\theta A}P_{BB}) + P_{AA}P_{AB} + P_{AB}P_{BB}\right] \quad (73J)$$

Recall that in the system studied in the main body of this report, the differential equations for the covariance terms between the amplitude and the frequency and phase functions were linear first order equations whose solutions would tend toward zero. Thus in the steady state, those two covariance terms could be assumed to be zero, considerably simplifying the solution. We shall investigate whether this simplification applies to the present problem.

First consider equation (73j) for $\dot{P}_{AB}$. If we collect all the terms containing P_{AB} on the right hand side, we obtain

$$<\dot{P}_{AB}> = -(\cdots)P_{AB}-(1/2N)\left[(A^2+B^2)P_{\theta A}P_{\theta B}+BP_{AA}P_{\theta B}-AP_{\theta A}P_{BB}\right] \tag{74}$$

This is again a linear first order differential equation in P_{AB}, with an input driving term containing A, B, $P_{\theta A}$, $P_{\theta B}$, P_{AA} and P_{BB}. While its homogeneous solution decays to zero, its complementary solution depends on the driving inputs.

Taking a similar look at the pair of equations (73f) and (73h) for $P_{\theta A}$ and $P_{\omega A}$, they can be written

$$<\dot{P}_{\theta A}> = -(\cdots)P_{\theta A}+P_{\omega A}-(1/2N)\left[B(P_{\theta\theta}P_{AA}+P_{\theta A}^2)-AP_{\theta\theta}P_{AB}+P_{\theta B}P_{AB}\right] \tag{75a}$$

and

$$<\dot{P}_{\omega A}> = -(\cdots)P_{\omega A}-(\cdots)P_{\theta A}-(1/2N)\left[BP_{\theta\omega}P_{AA}-AP_{\theta\omega}P_{AB}+P_{\omega B}P_{AB}\right] \tag{75b}$$

Again this pair of coupled equations has a homogeneous solution which decays to zero, but also may have a complementary solution

driven by some other terms. A similar situation occurs for the two equations for $P_{\theta B}$ and $P_{\omega B}$.

On inspection of this set of five equations, it is seen that even if all five covariances P_{AB}, $P_{\theta A}$, $P_{\theta B}$, $P_{\omega A}$, and $P_{\omega B}$ were initially zero, there are driving terms in their differential equations which will excite some non-zero response. An example of such a term is the $(B/2N)(P_{\theta\theta}P_{AA}+P_{\theta A}{}^2)$ term in equation (75a). Since the second factor is clearly positive and non-zero, the covariance element $P_{\theta A}$ obviously will become non-zero whenever B (the sine component of the input signal) is non-zero. This is one unfortunate result of the non-observability of the state of the system. Physically, whenever B is non-zero, there is no way to distinguish between a change in the other signal component amplitude A or a true change in the phase θ. Thus the covariance between A and θ will be non-zero.

These five covariance terms are different from the remainder of the covariance matrix though, in that they all are driven by A and B which are zero-mean random variables. Thus while the covariance values are not actually zero, they only fluctuate about zero in response to the A and B state inputs. Thus over the long term, their steady state value averages to zero. Inclusion of these terms unquestionably adds to the complexity of the state estimator and, since they are dynamically changing parameters, forces the implementation of the differential equations for at least this

part of the P matrix. However since they are due largely to the over-specification of the state of the system, it is felt that they can be dropped from at least a first order approximation to the optimum estimator. The estimator can then be designed using only those P matrix coefficients which have a non-zero steady state value and do not require continuous computation.

On the basis of this assumption, the steady state values for the remaining terms of the P matrix may be found by dropping the five P matrix terms discussed above and setting the remainder of the derivitives to zero. This leads to the set of four equations

$$\langle \dot{P}_{\theta\theta} \rangle = 2P_{\theta\omega} - (A^2+B^2)P_{\theta\theta}^2/2N = 0 \tag{76a}$$

$$\langle \dot{P}_{\omega\omega} \rangle = \alpha^2 Q_\omega - 2\alpha P_{\omega\omega} - (A^2+B^2)P_{\theta\omega}^2/2N = 0 \tag{76b}$$

$$\langle \dot{P}_{\theta\omega} \rangle = P_{\omega\omega} - \alpha P_{\theta\omega} - (A^2+B^2)P_{\theta\theta}P_{\theta\omega}/2N = 0 \tag{76c}$$

$$\langle \dot{P}_{AA} \rangle = \langle \dot{P}_{BB} \rangle = \beta^2 Q_A - 2\beta P_{AA} - P_{AA}^2/2N = 0 \tag{76d}$$

where it is recognized that the equations for P_{AA} and P_{BB} are identical. Comparing these equations with equations (19) in the main text shows that they are identical in form, except for the replacement of $A^2/2N$ with $(A^2+B^2)/2N$ for the signal to noise ratio. Thus the solutions for the steady state frequency and phase

estimation errors obtained in the main text, and the conclusions reached regarding performance as a function of signal to noise ratio and other parameters, remain unchanged. The interpretation of the solution for the amplitude estimation error must be changed slightly because Q_A is now related to the total signal power rather than to an amplitude modulation.

The next step in finding the form of the optimum estimator is to substitute the matrix definitions into the estimation equation

$$\dot{\hat{X}} = F \hat{X} + G w_o + P H^T R^{-1} (z - h(\hat{X}))$$

$$= \begin{bmatrix} \hat{\omega} \\ -\alpha\hat{\omega} \\ -\beta\hat{A} \\ -\beta\hat{B} \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha\omega_o \\ 0 \\ 0 \end{bmatrix} + (1/N)(z - \hat{A}\cos\hat{\theta} - \hat{B}\sin\hat{\theta}) \cdot$$

$$\begin{bmatrix} P_{\theta\theta}(B\cos\theta - A\sin\theta) + P_{\theta A}\cos\theta + P_{\theta B}\sin\theta \\ P_{\theta\omega}(B\cos\theta - A\sin\theta) + P_{\omega A}\cos\theta + P_{\omega B}\sin\theta \\ P_{\theta A}(B\cos\theta - A\sin\theta) + P_{AA}\cos\theta + P_{AB}\sin\theta \\ P_{\theta B}(B\cos\theta - A\sin\theta) + P_{AB}\cos\theta + P_{BB}\sin\theta \end{bmatrix} \tag{77}$$

This is the set of differential equations which must be solved if all P matrix values are assumed to be non-zero. However, if we retain only the five terms $P_{\theta\theta}$, $P_{\omega\omega}$, $P_{\theta\omega}$, P_{AA}, and P_{BB} which

have non-zero steady state values, and ignore the remaining covariance terms which fluctuate around zero, these equations simplify to

$$
\begin{bmatrix} \dot{\theta} \\ \dot{\omega} \\ \dot{A} \\ \dot{B} \end{bmatrix} = \begin{bmatrix} \omega \\ \alpha(\omega_o - \omega) \\ -\beta A \\ -\beta B \end{bmatrix} + (1/N)(z - A\cos\theta - B\sin\theta) \cdot \begin{bmatrix} P_{\theta\theta}(B\cos\theta - A\sin\theta) \\ P_{\theta\omega}(B\cos\theta - A\sin\theta) \\ P_{AA}\cos\theta \\ P_{BB}\sin\theta \end{bmatrix} \tag{78}
$$

Finally, if we again take short time averages to eliminate the effects of terms at the carrier frequency or higher, we obtain

$$
<\dot{\hat{\theta}}> = \hat{\omega} + (P_{\theta\theta}/N)<(\hat{B}\cos\hat{\theta} - \hat{A}\sin\hat{\theta}) \cdot z> \tag{79a}
$$

$$
<\dot{\hat{\omega}}> = \alpha(\omega_o - \hat{\omega}) + (P_{\theta\omega}/N)<(\hat{B}\cos\hat{\theta} - \hat{A}\sin\hat{\theta}) \cdot z> \tag{79b}
$$

$$
<\dot{\hat{A}}> = -\beta\hat{A} + (P_{AA}/N)(<z \cdot \cos\hat{\theta}> - \hat{A}/2) \tag{79c}
$$

$$
<\dot{\hat{B}}> = -\beta\hat{B} + (P_{BB}/N)(<z \cdot \sin\hat{\theta}> - \hat{B}/2) \tag{79d}
$$

Figure 11 shows this set of equations expressed in a block diagram form similar to that of Figure 1 in the main text. The upper part of the figure shows the estimators for the amplitude components A and B. Each is a simple low-pass filter averaging the mean product of the input signal with either the sine or the cosine phase output of the reference function generator. The time
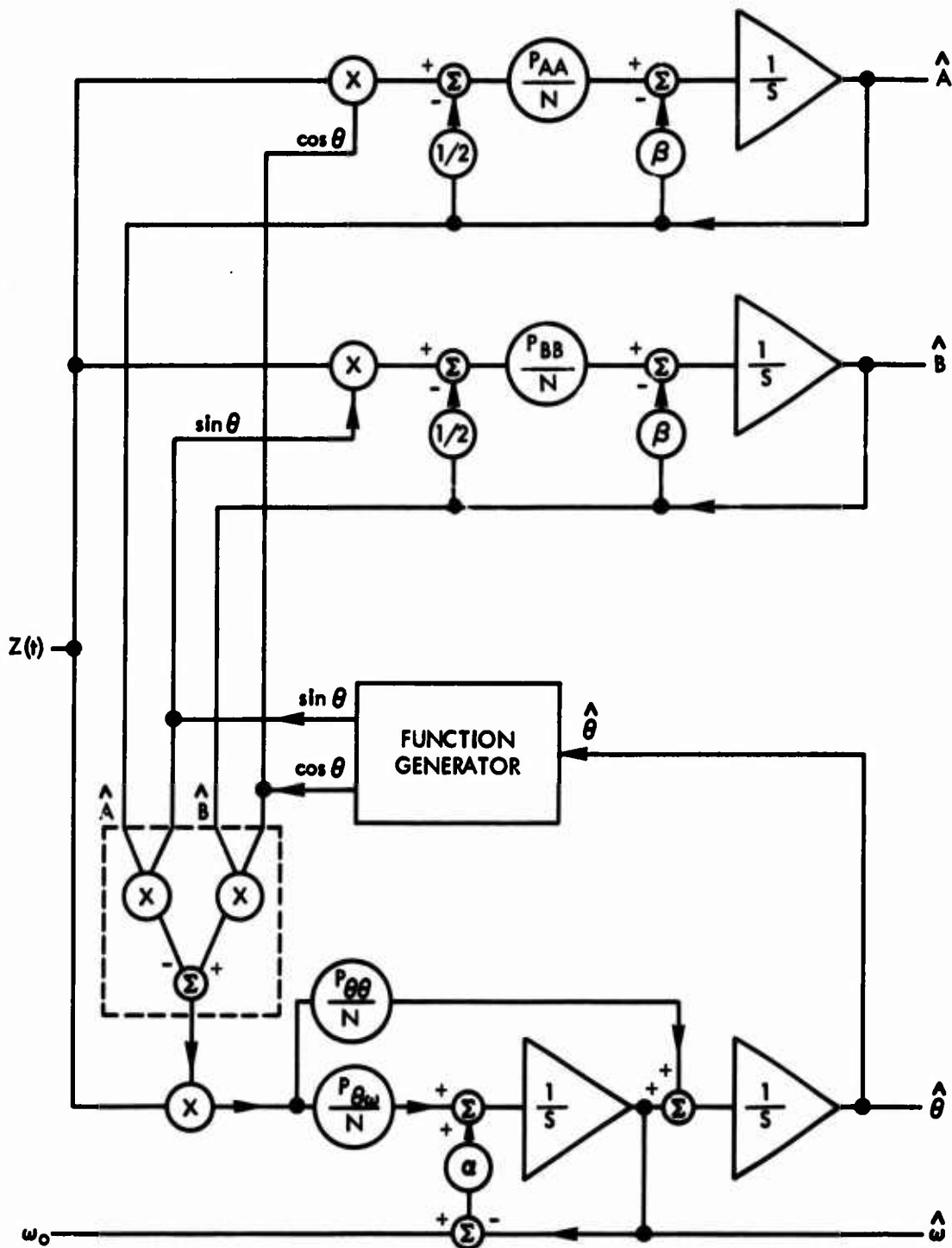
FIG. 11   FREQUENCY TRACKER STRUCTURE FOR RAYLEIGH FADING CASE

constant of this filter is determined by the value of the
uncertainty matrix element P_{AA} or P_{BB}, just as in the system
studied in the main text. In fact the only change due to the
Rayleigh fading model is that two channels of amplitude estimator
are required rather than just one.

The frequency and phase estimation portion of the circuit
is also quite like that of Figure 1 in the main text, in that
it is a second-order phase lock loop whose dynamics are determined
by the coefficients $P_{\theta\theta}$ and $P_{\theta\omega}$. Since these are determined by the
same equations in both of the models, the performance of the phase
lock loop should be identical.

The only real difference caused by the Rayleigh fading model
is in the reference function used by the phase lock loop. In the
simpler case, the PLL reference came directly from one phase
of the function generator output. In the Rayleigh case, the
function $B\cos\theta - A\sin\theta$ is used to form the reference signal for
the phase lock loop. Whenever A and B (or their estimates) are
constant, this difference is immaterial and the Rayleigh solution
degenerates to that of the previous case. However, when A and
B fluctuate as assumed in the Rayleigh model and these fluctuations
are tracked by the A and B estimators, the complex reference
function automatically changes its phase to compensate for the
new ratio of A and B. Thus phase changes caused by fading of
the signal phase components are handled by the amplitude estimation

network and do not require a response from the phase and
frequency estimation circuit. This is particularly important
in the case of deep signal fades which cause a rapid reversal in
the received signal phase, since it is these rapid phase changes
which are the most difficult for the phase lock loop to handle.

Obviously things are not quite this simple, again because
of the non-observability of the state. Whenever a change is
observed in the received signal phase, it can be interpreted
as either a frequency change or as a change in the signal
vector components. The degree to which this is handled by
the amplitude or the phase portion of the estimator system is
determined by the values α, β, and Q_ω used in modeling the signal
behavior. The fact that both portions of the system attempt to
respond to each change is the source of the fluctuating covariance
terms such as $P_{\theta A}$ which we chose to ignore in developing this
estimator.

The conclusion to be reached from this Appendix is that,
at least to a first approximation, the optimum frequency estimator
in the presence of Rayleigh fading is still a phase lock loop
system. It is somewhat more complex than in the pure amplitude
fading model in that two amplitude estimators are formed from
quadrature phases of the reference signal, and these estimators
are used to form the reference signal for the phase lock loop.
The primary effect of this change is that it assists the phase
lock loop in tracking the signal phase through deep fading periods.

REFERENCE

(1) Bryson, A.E. and Ho, Y.C., Applied Optimal Control,
     p 364, Blaisdell Publishing Company, 1969

GLOSSARY

A	signal amplitude
A_o	mean amplitude of signal
α	bandwidth of frequency fluctuations
β	bandwidth of amplitude fluctuations
B	amplitude of quadrature signal component
BW	bandwidth of the phase lock loop
D	rms frequency deviation in radians/sec
$\delta(\tau)$	Dirac delta function
F	matrix describing evolution of signal state
G	matrix coupling random process to signal state
G_A	input signal weight in amplitude estimator
H	linearized measurement matrix
h(X)	scalar measure of signal state
I_A	amplitude modulation index
I_ω	frequency modulation index $= D/\alpha$
K	a parameter of the phase lock loop
N	noise power per Hertz (double sided)
P	covariance matrix of estimation error
Q	covariance matrix of W
R	noise autocorrelation matrix
ρ	a signal to noise density ratio parameter
ρ_A	normalized signal to noise ratio $= \rho/\beta$
ρ_ω	normalized signal to noise ratio $= \rho/\alpha$
SNR	signal to noise ratio in 1 Hertz band (single sided)
T(s)	transfer function of phase lock loop

GLOSSARY (Cont.)

τ_A	time constant of amplitude estimator
θ	signal phase angle in radians
v	additive noise process
w	random process producing fluctuations in state
w_o	mean value of W
ω	signal frequency in radians/sec
ω_o	mean frequency of signal
X	signal state vector
$\dot{X}$	derivative of signal state vector
$\hat{X}$	estimate of signal state vector